Title
Essays on the Efficiency of Online Lending

Author
Chen, Baizhu

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

ESSAYS ON THE EFFICIENCY OF ONLINE LENDING

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ECONOMICS

by

Baizhu Chen

June 2018

The Dissertation of Baizhu Chen is approved:

Professor Nirvikar Singh, chair

Professor Georege Bulman

Professor Daniel Friedman

Professor Tyrus Miller
Vice Provost and Dean of Graduate Studies

Contents

List of Tables

List of Figures

Abstract

Essays on Efficiency of Online Lending

Baizhu Chen

This dissertation has three chapters, with emphasis on the efficiency of online lending using multiple identification strategies. The first chapter presents a theoretical model (1) to help illustrate the behavior of lenders and borrowers in a P2P market and (2) to derive a reduced form model for empirical analysis. In the model, borrowers decide whether or not to default. The benefit of default is that a borrower gets to increase consumption if he chooses not to repay. On the other hand, I assume that default imposes a direct utility loss on the borrower. This loss is assumed to be random and unobservable by lenders or anyone other than the borrower. *Ex ante*, when lenders choose whether or not fund a loan, they compare the expected return of the loan vs. their opportunity costs. The opportunity cost is also assumed to be random and unobservable by others. Also, the expected return of the loan depends on lender's subjective expectations of default risk. The subjective expectations are formulated based on the observed borrower characteristics and loan information. I present methods of testing whether the lenders' subjective expectations of default is the same as the objective ones.

In the second chapter, using a unique dataset of peer-to-peer (P2P) lending with detailed loan and borrower information, I study which borrower characteristics lenders value when choosing loans to fund, and whether lenders value characteristics that minimize the probability of default. In this online context, the researcher observes everything that the lender does, enabling unbiased estimation of borrower characteristics that lenders favor. However, estimating characteristics that predict loan default is problematic due to selection at the funding stage. I implement three

strategies to address this issue: (1) restricting attention to borrower characteristics for which there is no evidence of selection in the first stage; (2) exploiting variation in the probability of funding caused by contemporaneous competition on the platform; and (3) bounding the default estimates in the style of Lee (2009). The results imply that P2P lenders consider employment borrower characteristics as most important when making funding decisions, but disregard several characteristics that are valuable predictors of default risk. Specifically, lenders overestimate the importance of verified employment information, and underestimate the importance of verified education level and marital status.

The third chapter estimates the effect of credit insurance in the peer-to-peer (P2P) lending market. Online lending is a fast growing area of finance. However, it is often plagued by asymmetric information problems and high credit risk, especially in developing countries that do not have well-established financial and credit rating systems. To address this issue, some Chinese P2P marketplaces have incorporated loan insurance into their online platforms. We estimate the treatment effects of credit insurance on the P2P market by exploiting a unique quasi-experiment. Specifically, loan guarantees gradually became available on a top Chinese P2P lending platform to borrowers from 31 major cities between 2012 and 2014 through 12 waves of business expansion. Our empirical results suggest that the availability of credit insurance resulted in significant, strong, and persistent treatment effects on the market demand and supply. The adoption of credit insurance was associated with dramatic increases in the number of loan listings, funding probability per loan, and bidding amount per lender. The average funding time per loan decreased by about 170 hours.

Acknowledgements

I wish to express my sincere appreciation to the people who have provided me with great help and support in the completion of my dissertation and the pursuit of my Ph.D.

First and foremost I would like to thank my advisors Professor Nirvikar Singh, Professor George Bulman, and Professor Dan Friedman. It has been a great honor for me to be the Ph.D. student of three wonderful professors. It is such a fortune to have them on my dissertation committee for their unwavering guidance on my research and prompt responses to my questions and needs.

I want to thank Professor Singh for all the encouragement and enlightenment that he has given to me in exploring new ideas, valuable datasets, and different methodologies. His charming and optimistic personality inspired me to overcome difficulties and stay positive even at the toughest moments during my Ph.D. studies. In addition, Professor Singh was very generous in sharing his knowledge and network to help me start a successful career.

Also, I would like to extend my gratitude to Professor Bulman, who has been tirelessly advising my dissertation work in the last three years. I appreciate all his contributions of time, effort, and ideas that made my research experience productive and intriguing. I feel very grateful for what I have learned from the excellent example of Professor Bulman as a promising young economist and professor.

I want to thank Professor Friedman for his tremendous help in my research as well. Every talk I had with him was so enlightening. He introduced me to many people in academia and industry, from whom I got to learn about the research frontier in social networks, fin-tech, and Internet economics, broadening my mindset and vision.

Once again, I would like to thank my dissertation committee for their instruction and commitment. My gratitude for their contribution to my future success is immeasurable.

Furthermore, I would like to thank the other three members of my oral defense committee, Professor Michael Hutchison, Professor Eric Aldrich, and Professor Amilcar Menichini for their time and valuable comments. Also, I would like to thank all the participants in the applied micro workshops and the job market practice talks for their helpful comments and suggestions.

I need to thank the economics department and the CAFIN of UC Santa Cruz for the physical and financial support for my study and research. I want to thank Professor Jonathon Robinson for his leadership of the Ph.D. program and Sandra Reebie for her hard work and assistance. I would like to thank all the people in our Ph.D. program for their immense contribution to my personal and professional time at UC Santa Cruz.

My time at Santa Cruz was made so much enjoyable in large part due to the many friends that became a part of my life. I am grateful for the time spent with my friends and for many other people and memories. The friendships that we have built are so deep that they shall last forever.

Last but not least, I would like to thank my family for all their love and encouragement. Their persistent financial and, even more importantly, mental support is the key to my success. A special thank you goes to my mother who raised me with a love of scientific research and supported me in all my pursuits whole-heartedly. And most of all for my loving, encouraging, and patient boyfriend Wei who has given tremendous support during the last two years of my Ph.D. study. Without his help, this would not have been possible. Thank you so much.

Chapter 1

Peer-to-Peer Lending: A Theoretical Model and Its Empirical Implications

1.1 Introduction

The past two decades have witnessed remarkable growth in the digital economy thanks to the development of the Internet and the information exchange and networking it enables. Since the late 1990s, the online economy has quickly expanded from e-commerce – i..e., retail/wholesale businesses – into many other sectors, notably FinTech – a new industry that combines finance and information technology. After several years of rapid growth, FinTech, now consisting of several key segments such as digital payments, personal finance, and alternative finance, is redefining the financial sector.

In this chapter, I study the behavior of agents in alternative finance, which refers to financial channels, processes, and instruments outside of the traditional finance system such as banks and capital markets. In particular, I develop a theoretical model characterizing the optimization problems for individual borrowers and lenders on peer-to-peer (P2P) lending marketplaces. Being an alternative as well as a complement to conventional financing, P2P lending is the practice of borrowing and lending between unrelated individuals (or business entities) through online platforms instead of traditional channels such as banks and other financial institutions. It is the largest market segment in alternative finance, greatly exceeding crowdfunding, balance sheet lending, and invoice trading. The theoretical model derives regression equations that can be deployed to estimate impacts of different factors on *ex ante* funding decisions of lenders as well as *ex post* default decisions of borrowers with testable empirical implications. I implement the empirical approach in the second chapter in order to identify the most important factors that determine funding and default risk, and tp examine whether or not ex ante funding decisions are efficient.

P2P loans are primarily funded by a number of small investors, allowing an analysis of the investment behavior of individual lenders in the credit market. By comparison, the traditional finance literature mainly focuses on sophisticated investors, such as large banks. Additionally, P2P lending transactions are transparent to borrowers, lenders and the researcher, unlike traditional bank loans that frequently depend on borrower characteristics and human interactions that are unobserved by the researcher. Potential lenders can view a limited set of borrower and loan characteristics on which they make investment decisions. In the context of P2P lending, all the information about a loan request available to potential lenders is equally visible to the researcher. Utilizing all of these features to analyze behavior of individual borrowers and lenders, I build up a simple theoretical model in which lenders make funding decision to maximize their expected payoff conditional on the information set of all the observable borrower and loan characteristics, which partially though not perfectly determines a borrower's potential default risk. The derived empirical implication suggests that an efficient lender should value a borrower's characteristic more *ex ante* if it associates with lower *ex post* default risk, which is testable by a comparison of the regression coefficient of the characteristic in the funding stage and its counterpart in the default stage.

In addition, I discuss potential bias issues due to sample selection and omitted confounding variables when the model is taken to empirics with real world data. An omitted variable problem arises as the observed borrower and loan characteristics can be correlated with unobserved ones. However, the model suggests that this is less of a concern because lenders only care about the associations between observed characteristics and default. On the other hand, the sample selection issue poses a threatening challenge to the legitimacy of the empirical analysis. Because loans are not funded randomly and default records are available for funded loans only, it becomes invalid to compare directly between coefficients in the funding regression and those in the default regression as funded loans may not be comparable with unfunded loans. Two potential solution methods are discussed: Heckman selection model with an exogenous instrument for funding decision following Heckman (1979) and bounded treatment effects following Lee (2009).

1.2 Growth in Peer-to-Peer Lending

The global market for alternative finance has grown dramatically as documented in a series of industry reports composed by the University of Cambridge. According to the reports, the Asia Pacific region has been the largest market segment for alternative finance, with total market volume reaching $245.28 billion in 2016, an annual growth of 136% from $103.31 billion in the previous year. Within the region, $243.28 billion – nearly all of the regional volume – was raised in mainland China alone, making it the world's largest market for alternative finance. The runner-up position went to the US market and its market volume went up 22% to $34.5 billion in 2016 while the regional market volume of the Americas was $35.2 billion, a 23% year-on-year increase from 2015. The total European alternative finance market grew by 41% to $8.4 billion (€ 7.7 billion) in 2016, with the UK being the largest market in the region contributing 73% of the regional volume. More than three quarters of alternative financing was facilitated through online P2P lending marketplaces whose total market volume reached over $225 billion in 2016. The dominance of P2P lending in alternative finance is largely driven by the overwhelming volume of P2P lending in the Chinese and US markets. The dominance of P2P lending has decayed somewhat due to new business models such as crowdfunding and balance sheet lending that are drawing more and more popularity.[1]

The history of P2P lending dates back to the inception of Zopa – the world's first and the current largest European P2P lending platform – in the UK in February 2005, which was followed by the launching of the US marketplaces such as Prosper and Lending Club a year later. The growth in the P2P lending market has been dramatic since its genesis and speed up further during the Great Recession. Haliassos (2013) argues that heavy losses from the subprime mortgage crisis and tightened regulations forced many commercial banks to scale back lending

[1] See the series of alternative finance industry reports: *2017 The Americas Alternative Finance Industry Report*, *The 2nd Asia Pacific Region Alternative Finance Industry Report*, and *The 3rd European Alternative Finance Industry Report* for reference.

to small business owners and individual consumers, creating greater demand for alternative financing channels. Demyanyk and Kolliner (2014), a report published by the Cleveland Federal Reserve on P2P lending in the US, note that the total volume of P2P lending has been growing at an astonishing 84% per quarter, while the total amount of bank-originated consumer-finance loans and the total amount of bank-originated credit-card lending have been declining by an average of 2% and 0.7% per quarter respectively during the time period between 2007:Q2 and 2014:Q1. It was projected by PricewaterhouseCoopers that the total market volume in the US is going to grow from approximately \$5.5 billion in 2014 to \$150 billion or higher by 2025.[2] While Srethapramote et al. 2015, Morgan Stanley predict that the global market size of P2P will have an expected annual growth rate of 51% until at least 2020. On December 10, 2014, the San Francisco based market leader Lending Club (NYSE: LC) launched its IPO on New York Stock Exchange, the first within the industry, resulting in a market capitalization of over \$6 billion.

The idea of P2P lending is to redefine the roles of financial intermediaries by bringing together credit demand and supply through the Internet. Compared to the traditional financing channels such as bank loans and credit card lending, P2P lending has gained popularity among borrowers due to its lower entry barrier, easier application process, quicker funding decisions, and better accessibility. As pointed out in Demyanyk and Kolliner (2014), the rapid growth of P2P lending is foremost attributable to the remarkable contribution that it brought to the financial market: a substantial improvement in the access to credit in particular for individuals and small- and medium- sized businesses who are short of sufficient credit histories and collateral assets and hence are often ignored by traditional financial channels. They also note that credit from P2P lending is often cheaper than traditional personal loans, as average P2P interest rates (on Lending Club) have been lower than credit card rates since the first quarter in 2010, while the performance of P2P loans is comparable to that of bank-originated consumer loans. Individuals and SMEs constitute the majority of those who suffer credit rationing in traditional financial markets. As analyzed in the groundbreaking work of Stiglitz and Weiss (1981; 1983), when there exists imperfect information in the credit market, it is optimal for banks to adopt credit rationing in equilibrium instead of raising interest rates to clear the market, because higher interest rates would discourage safer borrowers (adverse selection) and induce borrowers to take

[2] See "Peer Pressure: How peer-to-Peer lending platforms are transforming the consumer lending industry" by PwC, February 2015 for reference .

riskier actions (moral hazard), and hence reduce the expected return to lenders. They note that it does not necessarily help correct the the market failure when banks enrich their strategy by including collateral requirements (Stiglitz and Weiss, 1992).

On the other hand, the emergence of P2P lending attracts many individual investors who possess limited wealth and often have few investment opportunities. Because an individual investor does not have to finance an entire loan request, it lowers the barrier to become a P2P lender and has the potential to improve investment diversification. As a leading segment in FinTech, P2P lending is particularly popular among younger generations who are more familiar with the Internet and digital technologies, and may be more resentful of Wall Street and the traditional financial sector. Recently, online lending started drawing the attention of institutional investors.

1.3 Literature Review

This chapter contributes to a large literature on the determinants of loan funding and loan performances. This literature mostly focuses on impacts of information asymmetry on credit supply, cost of funding, and risk sharing. Intuitively, information sharing is among the most crucial factors due to its direct effect on resolving information asymmetry between borrowers and lenders. Jappelli and Pagano (2002) theoretically predict that information sharing among lenders attenuates adverse selection and moral hazard, and can therefore increase lending and reduce default risks, a prediction that is supported by cross-country empirical evidence. Djankov, Mcliesh and Shleifer (2007) investigate the determinants of private credit in over 100 countries, and find that improvements in creditor rights and in information sharing leads to increases in private credit. Schenone (2010) finds that information sharing can result in information rents in the credit market. However, the effect of information sharing may depend on the level of social trust as pointed out in Pevzner, Xie and Xin (2015). Analyzing the traditional bank loan funding process, Kim, Surroca and Tribo (2014) find that cultural proximity (e.g., shared codes, beliefs, ethnicity) between lenders and borrowers contributes to lower information frictions in lending, which is echoed in Fisman, Paravisini and Vig (2017). Jiang, Nelson and Vytlacil (2014) find that borrower and loan characteristics and the authenticity of the information explain a large fraction of mortgage delinquency rates. When less credit-worthy borrowers are ignored by formal financial institutions and forced to borrow from the informal market instead, Bose (1998) shows that providing cheap credit through the formal sector can generate adverse 'composition effects' which worsen the terms of credit and the availability of loans in the informal one due to information asymmetry between lenders and borrowers in the informal sector. Morrison (2005) models the effect of introduction of reporting requirements for credit derivatives and his results illustrate the reporting requirements can effectively prevent welfare reduction. When information asymmetry prevails in the credit market, rational outside investors will try to infer the insiders' information from the firm's financial structure. However, Flannery (1986) shows in the absence of market transaction cost a firm's financial structure cannot provide a valid signal, while the existence of a signaling equilibrium depends on the distribution of firms' quality and the magnitude of underwriting costs for corporate debt.

Another frequently-adopted method to overcome information asymmetry are collateral requirements. Inderst and Mueller (2007) argue that collateral mitigates the inefficiency in credit supply in an imperfectly competitive loan market with information asymmetry despite the fact that collateralized loans are more likely to default *ex post* once controlling for borrower's observable risk. Niinimaki (2009) shows that fluctuating value of loan collateral can generate a problem of moral hazard as banks intend to finance risky projects against collateral and rely on the rising collateral value to earn profits. Reputation (or repeated interactions/games between creditors and borrowers) is also proposed as a remedy for asymmetric information. Greenbaum, Kanatas and Venezia (1989) show lenders can derive informational advantages from the durability of information acquired as a result of an extant relationship with a client. However, Sharpe (1990) argues that customer relationships generate a different type of information asymmetry by allowing a bank to learn more than others about its own customers. This will result in market inefficiency when the allocation of capital shifts toward lower quality and inexperienced firms as competition drives banks to lend to new firms at low interest rates in the hope of capturing information rents in the future.

The transparency of P2P lending frees researchers from concerns of omitted variables in creditors' optimization decision making. So, different from the existing literature that specifically models and analyzes information asymmetry between borrowers and creditors, the model presented here is constructed mainly from the standing point of a lender by simplifying interactions between borrowers and lenders. Instead of focusing on either *ex ante* funding or *ex post* repayment, the model implication shows that one can empirically test whether lenders's funding decisions are efficient by analyzing the two phrases together.

1.4 Borrowing and Lending on P2P Marketplaces

Distinct from traditional financial intermediaries, P2P platforms do not receive funds from investors or issue loans to borrowers directly. Instead, they set up online marketplaces where borrowers get to post their loan requests among which lenders choose to invest. To facilitate loans, most platforms provide services as follows and generate revenues by charging service fees to borrowers and lenders. P2P lending platforms are typically responsible for the operations as follows: (a) marketplace (i.e., websites or webpages) construction and operation, (b) information verification (for borrowers and lenders) , (c) record keeping, (d) credit evaluation and risk control, (e) processing transactions and loan servicing, (f) legal governance and compliance, and (g) marketing.

Figure 1.1 illustrates the borrowing procedure on a P2P lending platform. To apply for a loan, a borrower needs to submit a loan application to the P2P lending platform specifying his borrowing purpose as well as the amount and term requested. In most cases, together with the application, the borrower needs to present required information on personal and financial conditions as well as copies of related documents and proofs to the platform who will then assess the loan request and the qualifications of the borrower. Once the application is approved by the platform, the loan request will be listed on the online marketplace with all the available borrower and loan characteristics to solicit interested lenders. Interested investors can bid to fund the borrower by buying notes of the loan (often on a first-come-first-serve basis).[3] If the loan is 100% funded within the funding period, it becomes successfully funded and funds will be transferred to the borrower soon afterwards; otherwise, the loan request fails. In the coming month, the borrower will start repaying the loan with monthly installments.

Most P2P platforms posit a maximum amount of a single loan that a borrower can request and restrict the total number (and total amount) of loans that a borrower can owe through the marketplace. These constraints are often customized conditional on an individual's borrowing

[3] A note represents a small fraction of a loan and claim on its repayments. Its value equals the minimum investment per loan per investor set by each P2P platform. For instance, the minimum investment per loan is $25 on Prosper and Lending Club in the US and about $8 (50 yuan) on many Chinese platforms. This allows even investors with little wealth to invest in P2P loans.

history. With a longer and better record of borrowing and repayment, a borrower can have a higher cap on total borrowing through the platform. Some platforms allow borrowers to choose the length of borrowing terms from as short as couple of months to as long as several years; while others apply a specific term to all loans.[4] Determination of loan interest rate varies between platforms. Typically, there are three different methods of setting the rate: (1) using an auction mode,[5] (2) a fixed rate based on the borrower's risk level, and (3) a rate chosen by the borrower. Large platforms usually provide service for a general population with different kinds of borrowing purposes; while small ones may specialize in specific submarkets such as student loans, auto loans, real estate loans, etc. Nonetheless, they all generate revenues by charging borrowers (or lenders) small fees for facilitating and servicing loans and repayments.

In the model, I don't take a stand on how loan terms and interest rates are determined specifically and I allow borrowers to present a variety of personal characteristics both verifiable and non-verifiable. Because the design of the model shows how one can empirically identify a lender's efficiency, loan characteristics as well as borrower characteristics are assumed to be given exogenously from the point view of lenders, though they can correlate endogenously with default risk.

1.5 Theoretical Model

To answer the question of whether the lenders make optimal funding decisions in the P2P market with respect to borrower characteristics and information verification, I first represent the decision making of a typical borrower and a typical lender in a parsimonious model. It derives the regression models linking the underlying parameters to the empirical one. This also paves the way for the discussion in Section 1.6 about omitted variable problems, testable implications and sample selection issues. At the beginning of period t, N_t borrowers enter the market and apply for loans at the P2P platform. The number of the borrowers, N_t, is assumed to be exogenous and it varies over time. When applying for a loan in period t, borrower i posts a one-period loan

[4] Most platforms allow borrowers to repay their borrowing back fully in advance with a small or no penalty fee.

[5] For example, Prosper 1.0 (the first iteration of Prosper from 2006 to 2008 before it obtained SEC registration) used a Dutch-auction-like model to determine loan rates: a lender bids an amount and an interest rate that she is willing to lend, while the actual loan rate is determined by the lowest interest rate(s).

request of principal b_{it} and net interest rate r_{it}, where $b_{it} \in [b_{\min}, b_{\max}]$ and $r_{it} \in [r_{\min}, r_{\max}]$.[6] There are K verifiable borrower characteristics (e.g., ID, education, marital status, etc.) which borrower i can choose to verify in order to earn the trust of potential lenders and to increase the probability of being successfully funded. Let's denote VI_{kit} as the dummy variable that indicates whether borrower i has verified information k or not, and $\mathbf{VI}_{it} \equiv [VI_{1it}, \ldots, VI_{Kit}]^T$ as the column vector consisting of all K verification indicators. Verification characters are determined independently by the individual borrowers and are taken as a given in the following analysis because the research focuses on the funding decisions of lenders.

If the loan request is successful, borrower i receives the funds by the end of period t. He either repays or defaults on the loan at the beginning of the following period $t+1$, and N_{t+1} new borrowers enter the market. In this empirical part, borrowers of different periods are pooled together where repeated borrowers are taken as new borrowers. In reality, some P2P platforms allow borrowers to borrow for terms of different lengths in months while others specify a term for all loans. The assumption that the P2P borrowers borrow for one-period is made to simplify the model because loan term optimization is not the focus of the analysis. Loan term should be included as a control variable in the empirical analysis.

1.5.1 Borrower's Default Decision

Suppose at period t, borrower i's loan request is successfully funded. For simplicity, let's assume that the borrower's default decision is a discrete problem of choice between complete default and fully repaying the loan. Borrower i chooses to default if and only if the cost of full repayment is higher than the cost of default.

If borrower i chooses to pay back the loan, repayment occurs at the beginning of the next period and borrower utility is given by

$$U_{it}^{Repay} = U\left(w_{it} - m_{it}\right),$$

where w_{it} denotes borrower i's wealth level at the end of period t, $m_{it} = (1 + r_{it})\, b_{it}$ the full loan liability (the principal and the interests accrued) that borrower i has to pay back, and $U(\cdot)$ is a concave utility function. The utility level increases in the borrower's net wealth level $w_{it} - m_{it}$.

[6] Without loss of generality, b and r are assumed to be continuous variables and have independent boundaries. The subscript it is non-separable and it indicates borrower i who posts a loan request at period t.

When borrower i posts a loan request on the P2P platform, the loan contract (i.e., b_{it} and r_{it}) and hence the full loan liability is determined and known to all.

However, the current and potential wealth level is unobserved by lenders, resulting in asymmetric information. Borrower i knows the value of w_{it} when making the default decision, but creditors do not. Intuitively, the wealth level decreases and makes it costlier to fully repay the loan when a borrower draws an undesirable idiosyncratic shock (e.g., a negative liquidity shock) or suffers from a detrimental aggregate shock (e.g., becoming unemployed during an economic recession). The idiosyncratic shocks constitute private information of the borrowers and are known only to the borrowers themselves. On the other hand, the aggregate shocks are likely to be public information and known to both borrowers and the lenders, though the specific impacts may be different among the individual borrowers. Therefore, it is possible that the wealth levels are partially correlated across the borrowers within the same period.[7]

On the other hand, if borrower i chooses to default, there are two possible outcomes. With probability $P(\mathbf{VI}_{it})$, the borrower gets to walk away without paying (i.e., a "successful" default) and with probability $1 - P(\mathbf{VI}_{it})$, the borrower is sued by the P2P platform and has to repay the loan fully (i.e., a "failed" default).[8] In addition, the default imposes a random utility loss ψ_{it} on borrower i. The exact value of ψ_{it} is borrower i's private information and is not observed by the others. The introduction of the direct utility loss caused by default captures the idea that borrowers differ with respect to their credibility levels, and a more credible borrower has a higher value of ψ and hence, all else equal, lower utility from default. Therefore, borrower i's expected utility of default is given by

$$U_{it}^{Default} = \left[1 - P(\mathbf{VI}_{it})\right] U(w_{it} - m_{it}) + P(\mathbf{VI}_{it}) U(w_{it}) - \psi_{it}.$$

The probability of a "successful" default $P(\cdot)$ is assumed to be a weakly decreasing function in each element in the column vector $\mathbf{VI}_{it}$, i.e., $P_k \equiv \partial P / \partial VI_{kit} \leq 0$ for all $k \in \{1, 2, \ldots, K\}$. Consequently, the expected utility of default decreases weakly if borrower i has verified more information. The underlying intuition is that when a borrower verifies genuine personal informa-

[7]In the empirical part, the cross-sectional correlation between the borrowers is controlled by including time fixed effects.

[8]Only strategic defaults are considered in the discussion here. The possibility of insolvency or bankruptcy is ignored and every P2P borrower is assumed to be capable of fully repaying his loan considering the fact that the amount of a typical P2P loan is relatively small compared to individual's asset holdings.

tion such as ID and working related information to the P2P platform and the potential lenders, it is relative easier to find the person and launch lawsuits against the borrower to reclaim the loan back in the case of default.

Notice that although w_{it} and ψ_{it} are not directly observed by people other than borrower i, both variables are potentially correlated with the borrower's choice of information verification $\mathbf{VI}_{it}$. This is because, when applying for a loan, a borrower who is unlikely to be constrained by a future negative liquidity shock is likely to verify more personal information (i.e., w_{it} and VI_{kit} are potentially positively correlated) while a dishonest borrower is less likely to do so (i.e., ψ_{it} and VI_{kit} are positively correlated).

Consequently, borrower i's discrete choice of default is determined by the following decision rule. Borrower i chooses to default (i.e., $D_{it} = 1$) if the expected utility of default is higher, and fully repays the loan (i.e., $D_{it} = 0$) otherwise. Specifically,

$$D_{it} = \begin{cases} 1 & \text{if } U_{it}^{Default} > U_{it}^{Repay}, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

Without making assumptions on the specific functional forms of $U(\cdot)$ and $P(\cdot)$, Equation (1.1) can be rephrased in a linearized form using the first-order approximation. Let's begin by representing the conditions in Equation (1.1) in the equivalent logarithm terms as follows

$$D_{it} = \begin{cases} 1 & \text{if } \ln P\left(\mathbf{VI}_{it}\right) + \ln\left[U\left(w_{it}\right) - U\left(w_{it} - m_{it}\right)\right] - \ln \psi_{it} > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

The first-order approximation of $\ln P\left(\mathbf{VI}_{it}\right) + \ln\left[U\left(w_{it}\right) - U\left(w_{it} - m_{it}\right)\right] - \ln \psi_{it}$ is

$$\begin{aligned} &\ln P\left(\mathbf{VI}_{it}\right) + \ln\left[U\left(w_{it}\right) - U\left(w_{it} - m_{it}\right)\right] - \ln \psi_{it} \\ \approx & C_1 + \sum_{k=1}^{K} P_k\left(\bar{\mathbf{VI}}\right)\left[P\left(\bar{\mathbf{VI}}\right)\right]^{-1} VI_{kit} + \left(\frac{U'\left(\bar{w}\right) - U'\left(\bar{w} - \bar{m}\right)}{U\left(\bar{w}\right) - U\left(\bar{w} - \bar{m}\right)}\right) w_{it} \\ &+ \left(\frac{U'\left(\bar{w} - \bar{m}\right)}{U\left(\bar{w}\right) - U\left(\bar{w} - \bar{m}\right)}\right) m_{it} - \frac{\psi_{it}}{\bar{\psi}}, \end{aligned} \tag{1.3}$$

where

$$
\begin{aligned}
C_1 = & 1 + \ln P\left(\bar{\mathbf{V}}\mathbf{I}\right) + \ln\left[U\left(\bar{w}\right) - U\left(\bar{w} - \bar{m}\right)\right] - \ln\bar{\psi} \\
& - \sum_{k=1}^{K} P_k\left(\bar{\mathbf{V}}\mathbf{I}\right)\left[P\left(\bar{\mathbf{V}}\mathbf{I}\right)\right]^{-1}\bar{V}I_k - \frac{U'\left(\bar{w}\right) - U'\left(\bar{w} - \bar{m}\right)}{U\left(\bar{w}\right) - U\left(\bar{w} - \bar{m}\right)}\bar{w} \\
& - \frac{U'\left(\bar{w} - \bar{m}\right)}{U\left(\bar{w}\right) - U\left(\bar{w} - \bar{m}\right)}\bar{m},
\end{aligned}
$$

$P_k \equiv \partial P / \partial V I_k$, and $\bar{x}$ denotes the mean of variable x. And because $m = (1 + r)\, b$, thus

$$
m \approx \bar{m} + \bar{b}\left(r - \bar{r}\right) + \left(1 + \bar{r}\right)\left(b - \bar{b}\right) \tag{1.4}
$$

Plugging Equation (1.4) and Equation (1.3) into Equation (1.2) gives the linearized default decision rule,

$$
D_{it} = \begin{cases} 1 & \text{if } C_d + \sum_{k=1}^{K} \beta_{1k} V I_{kit} + \beta_2 r_{it} + \beta_3 b_{it} + \beta_4 w_{it} + \beta_5 \psi_{it} > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{1.5}
$$

where

$$
\begin{aligned}
C_d &= C_1 - \frac{\bar{r}\bar{b}U'\left(\bar{w} - \bar{m}\right)}{U\left(\bar{w}\right) - U\left(\bar{w} - \bar{m}\right)}, \\
\beta_{1k} &= P_k\left(\bar{\mathbf{V}}\mathbf{I}\right)\left[P\left(\bar{\mathbf{V}}\mathbf{I}\right)\right]^{-1}, \quad \forall k \in \{1, 2, \ldots, K\} \\
\beta_2 &= \frac{\bar{b}U'\left(\bar{w} - \bar{m}\right)}{U\left(\bar{w}\right) - U\left(\bar{w} - \bar{m}\right)}, \\
\beta_3 &= \frac{\left(1 + \bar{r}\right)U'\left(\bar{w} - \bar{m}\right)}{U\left(\bar{w}\right) - U\left(\bar{w} - \bar{m}\right)}, \\
\beta_4 &= \frac{U'\left(\bar{w}\right) - U'\left(\bar{w} - \bar{m}\right)}{U\left(\bar{w}\right) - U\left(\bar{w} - \bar{m}\right)}, \\
\beta_5 &= -\frac{1}{\bar{\psi}}.
\end{aligned}
$$

where C and β's are constant coefficients. β_{1k} is the marginal effect of the probability of default if information k is verified. β_2 and β_3 measure the change in probability of default with respect to change in interest rate and loan amount. β_4 and β_5 are the effects of the unobservable w and ψ on loan default rate.

In the analysis that follows, I represent Equation (1.5) in the form of a linear probability model for the purpose of simplicity, i.e., the probability that borrower i defaults on his loan is determined by the following linear equation,

$$D_{it} = \beta_0 + \sum_{k=1}^{K} \beta_{1k} VI_{kit} + \beta_2 r_{it} + \beta_3 b_{it} + \beta_4 w_{it} + \beta_5 \psi_{it}. \tag{1.6}$$

Effects of Information Verification on the Default Probability

Equation (1.6) is the true linear probability model that determines the borrower's default decision conditional on all the related variables. Among these determinants of the loan default probability, the borrower information verification indicators, $\mathbf{VI}_{it}$, and loan contract information, r_{it} and b_{it}, are determined before the default decision is made and are observed by borrowers, potential lenders, and by the researcher. On the other hand, as illustrated before, w_{it} and ψ_{it} are random variables that are known to neither the lenders nor the researcher. And more importantly, both variables are potentially correlated with VI_{kit} for all $k \in \{1, 2, \dots, K\}$. Exploiting the lenders-observable information only, the regression of the default probability equation that omits w and ψ are shown in Equation (1.7),

$$D_{it} = \tilde{\beta}_0 + \sum_{k=1}^{K} \tilde{\beta}_{1k} VI_{kit} + \tilde{\beta}_2 r_{it} + \tilde{\beta}_3 b_{it} + e_{it}, \tag{1.7}$$

would produce biased estimates for the coefficients of interest $\{\beta_{1k}\}_{k\in\{1,2,\dots,K\}}$.[9] The relationship between the corresponding coefficients in Equation (1.6) and Equation (1.7) can be expressed as $\tilde{\beta}_{1k} = \beta_{1k} + \beta_4 \delta_{1k} + \beta_5 \lambda_{1k}$ where δ_{1k} and λ_{1k} are obtained by projecting w and ψ onto the linear space spanned by $\{VI_k\}_{k\in\{1,2,\dots,K\}}$ and $\{r, b\}$.

$$w_{it} = \delta_0 + \sum_{k=1}^{K} \delta_{1k} VI_{kit} + \delta_2 r_{it} + \delta_3 b_{it} + e_{wit}, \tag{1.8}$$

$$\psi_{it} = \lambda_0 + \sum_{k=1}^{K} \lambda_{1k} VI_{kit} + \lambda_2 r_{it} + \lambda_3 b_{it} + e_{\psi it}, \tag{1.9}$$

where e_w and e_ψ are assumed to be disturbances independent of VI_k for all $k \in \{1, 2, \dots, K\}$.

To get the relationships between coefficients in Equation (1.6) and Equation (1.7), let's start

[9] The estimates for the other coefficients are also potentially biased compared to their true values, but they are not the coefficients of interest in this research.

by plugging Equation (1.8) and (1.9) into Equation (1.6), which results in the following equation

$$
\begin{aligned}
D_{it} = & \beta_0 + \sum_{k=1}^{K} \beta_{1k} VI_{kit} + \beta_2 r_{it} + \beta_3 b_{it} \\
& + \beta_4 \left(\delta_0 + \sum_{k=1}^{K} \delta_{1k} VI_{kit} + \delta_2 r_{it} + \delta_3 b_{it} + e_{wit} \right) \\
& + \beta_5 \left(\lambda_0 + \sum_{k=1}^{K} \lambda_{1k} VI_{kit} + \lambda_2 r_{it} + \lambda_3 b_{it} + e_{\psi it} \right).
\end{aligned}
$$

Rearrange the equation leads to Equation (1.7) where

$$
\begin{aligned}
\tilde{\beta}_0 &= \beta_0 + \beta_4 \delta_0 + \beta_5 \lambda_0, \\
\tilde{\beta}_{1k} &= \beta_{1k} + \beta_4 \delta_{1k} + \beta_5 \lambda_{1k}, \quad \forall k \in \{1, 2, \ldots K\}, \\
\tilde{\beta}_2 &= \beta_2 + \beta_4 \delta_2 + \beta_5 \lambda_2, \\
\tilde{\beta}_3 &= \beta_3 + \beta_4 \delta_3 + \beta_5 \lambda_3,
\end{aligned}
$$

and $e_{it} = \beta_4 e_{wit} + \beta_5 e_{\psi it}$, which implies that the disturbance in Equation (1.7) is orthogonal to VI_k, and this allows the coefficients $\left\{ \tilde{\beta}_{1k} \right\}_{k \in \{1,2,\ldots,K\}}$ to be estimated using simple techniques such as the least squares method.

Though the coefficient estimates using Equation (1.7) are biased compared to their true values, from the perspective of a lender, Equation (1.7) is nearly as valuable because *ex ante* it has the same power as the true model to predict the loan default probability given the lenders-observable loan information set. The coefficient of interest in Equation (1.7), $\tilde{\beta}_{1k}$, captures: (i) the direct impact of verifying borrower information k on default probability, i.e., β_{1k}, and (ii) the indirect impactor of the unobserved w_{it} and ψ_{it} that project onto the verification choice of information k, i.e., $\beta_4 \delta_{1k}$ and $\beta_5 \lambda_{1k}$ respectively. This is what matters from the lender's perspective.[10] Accordingly, conditional on the lenders-observable loan information set defined by the column vector $\mathbf{X}_{it} \equiv \left[\mathbf{VI}_{it}^T, r_{it}, b_{it} \right]^T$, the *ex ante* objective expectation of borrower i's default probability is given by

$$
\Psi\left(\mathbf{X}_{it}\right) \equiv \mathbb{E}\left(D_{it} | \mathbf{VI}_{it}, r_{it}, b_{it}\right) = \tilde{\beta}_0 + \sum_{k=1}^{K} \tilde{\beta}_{1k} VI_{kit} + \tilde{\beta}_2 r_{it} + \tilde{\beta}_3 b_{it}. \tag{1.10}
$$

[10]See Section 1.6 for more illustration.

1.5.2 Lender's Funding Decision

At period t, there is a continuum of potential P2P lenders of mass L_t each of whom is endowed with 1 unit of indivisible capital to invest. Every potential lender is matched with one borrower at any period, and the lender decides whether or not to invest in the loan request of the borrower she is matched with.[11] The probability that a borrower is matched with loan i is given by n_{it}, where $\sum_{i=1}^{N_t} n_{it} = 1$. Accordingly, loan i is matched with $L_t n_{it}$ potential lenders and a lower n_{it} implies that loan i is matched with fewer potential lenders. The matching probability n_{it} is determined exogenously and it can be decomposed as follows,

$$n_{it} = \frac{1}{N_t} + \pi_{it},$$

where $\sum_{i=1}^{N_t} \pi_{it} = 0$. $\frac{1}{N_t}$ is the common factor that affects the matching probability for every loan posted at time t. The matching probability is lower on average for every loan when there are more loan requests in a period (i.e., when N_t is larger).[12] While π_{it} is the idiosyncratic factor that affects the matching probability for loan it, when π_{it} is higher, then the matching probability is higher for loan it.

Suppose lender j is matched with borrower i. When making the funding decision, lender j forms a subjective expectation of the potential default probability of loan i conditional on the lenders-observable loan information set $\mathbf{X}_{it}$ using the linear equation as follows,

$$\Theta\left(\mathbf{X}_{it}\right) = \theta_0 + \sum_{k=1}^{K} \theta_{1k} VI_{kit} + \theta_2 r_{it} + \theta_3 b_{it}. \tag{1.11}$$

The subjective expectation function $\Theta(\cdot)$ is assumed to have a linear functional form similar to the true formula, Equation (1.10), except that its coefficients are potentially different. The intuition for the possible difference in the coefficients is that the lenders may behave irrationally or they may lack sufficient information to form the correct expectation of a borrower's default

[11] In practice, a P2P lender can choose the amount that she would like to invest in a loan and can invest in multiple loans at a time. However, this does not contradict against the assumption made in the model. The investment amount is determined by the number of loan shares that a lender intends to purchase. So one unit of capital here can be considered as equivalent to one loan share, and a lender who lends out m units of capital is equivalent to m identical lenders each of whom invests 1 unit. Similarly, a lender who invests in n different loans is equivalent to k identical lenders each of whom invests in a different loan.

[12] In reality, the matching probability of a loan request is mostly determined by the timing when the request is posted online and its location on the platform's webpage which are random according to the policy of the P2P platform.

probability. Irrational behavior can arise if the lender's prior consists of discrimination against certain borrower characteristics and hence associates the borrowers of such characteristics with higher probability of default. Alternatively, the lender's subjective expectation can deviate from the objective one when they lack sufficient information to make correct estimation of the expectation coefficients because of costly information acquisition.

Suppose lender j invests in loan i and the loan is successfully funded, then she would receive $\zeta \leq 1$ fraction of the principal if borrower i defaults and the full repayment (i.e., the principal plus the interests accrued) in return otherwise. So lender j's perceived expected utility of funding loan i is

$$G_{it}^{Fund} = \left[1 - \Theta\left(\mathbf{X}_{it}\right)\right] G\left(1 + r_{it}\right) + \Theta\left(\mathbf{X}_{it}\right) G\left(\zeta\right),$$

where $G(\cdot)$ denotes the concave utility function of the lenders. Because the model assumes that all P2P lenders use the same subjective expectation equation, the perceived expected utility of funding loan i is the same among all the lenders.

Instead, if a lender j chooses not to invest in the loan, she purchase a fixed-income investment that generates a guaranteed return of $\xi_{jt} \geq 0$. ξ_{jt} is a stochastic variable with a cumulative distribution function $\Xi(\cdot)$. The realization of ξ_{jt} is known only to lender j before making an investment decision.[13]

Therefore, lender j's funding decision rule is given as follows. Lender j chooses to fund loan i (i.e., $F_{ijt} = 1$) if and only if she believes the return of loan i out performs the random opportunity cost, i.e.,

$$F_{ijt} = \begin{cases} 1 & \text{if } p_{it} G_{it}^{Fund} + \left(1 - p_{it}\right) G\left(1 + \xi_{jt}\right) \geq G\left(1 + \xi_{jt}\right), \\ 0 & \text{otherwise,} \end{cases} \tag{1.12}$$

where p_{it} is the probability that loan i is successfully funded. Without making assumptions on the specific functional form of $G(\cdot)$, the funding decision rule implies that investor j is willing to lend to borrower i at time t if and only if her outside option ξ_{jt} is low enough as stated in Proposition 1.1.

[13] In reality, when a lender purchases some shares of a P2P loan, she deposits the funds to the account of the platform instead of that of the borrower. The borrower will receive the loan funds from the platform if and only if his loan request is successfully funded within the funding period. Otherwise, the funds will be reimbursed back to the lenders.

Proposition 1.1. *For loan i posted at period t, there exists an opportunity cost threshold $\xi_{it} \equiv C_f + \sum_{k=1}^{K} \gamma_{1k} VI_{kit} + \gamma_2 r_{it} + \gamma_3 b_{it}$. Those matched lenders who invest in the loan must draw opportunity costs below the threshold of which the probability equals $\Xi\left(\xi_{it}\right)$.*

Proof. The lender's funding decision rule, Equation (1.12), is equivalent to

$$F_{ijt} = \begin{cases} 1 & \text{if } \left[1 - \Theta\left(\mathbf{X}_{it}\right)\right] G\left(1 + r_{it}\right) + \Theta\left(\mathbf{X}_{it}\right) G\left(\zeta\right) - G\left(1 + \xi_{jt}\right) > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{1.13}$$

The first-order approximation of $\left[1 - \Theta\left(\mathbf{X}_{it}\right)\right] G\left(1 + r_{it}\right) + \Theta\left(\mathbf{X}_{it}\right) G\left(\zeta\right) - G\left(1 + \xi_{jt}\right)$ is

$$\begin{aligned} & \left[1 - \Theta\left(\mathbf{X}_{it}\right)\right] G\left(1 + r_{it}\right) + \Theta\left(\mathbf{X}_{it}\right) G\left(\zeta\right) - G\left(1 + \xi_{jt}\right) \\ \approx & C_2 + \sum_{k=1}^{K} a_{1k} VI_{kit} + a_2 r_{it} + a_3 b_{it} - a_4 \xi_{jt} \end{aligned} \tag{1.14}$$

where

$$\begin{aligned} C_2 = & G\left(1 + \bar{r}\right) - G\left(1 + \bar{\xi}\right) \\ & - \left[1 - \Theta\left(\bar{\mathbf{X}}\right)\right] G'\left(1 + \bar{r}\right) \bar{r} + G'\left(1 + \bar{\xi}\right) \bar{\xi}, \\ a_{1k} = & \left[G\left(\zeta\right) - G\left(1 + \bar{r}\right)\right] \theta_{1k}, \quad \forall k \in \left\{1, 2, \ldots, K\right\}, \\ a_2 = & \left[1 - \Theta\left(\bar{\mathbf{X}}\right)\right] G'\left(1 + \bar{r}\right) + \left[G\left(\zeta\right) - G\left(1 + \bar{r}\right)\right] \theta_2, \\ a_3 = & \left[G\left(\zeta\right) - G\left(1 + \bar{r}\right)\right] \theta_3, \\ a_4 = & G'\left(1 + \bar{\xi}\right), \end{aligned}$$

and $\bar{x}$ denotes the mean of variable x. Plugging Equation (1.14) into Equation (1.13) gives the linearized funding decision rule Equation (1.15)

$$F_{ijt} = \begin{cases} 1 & \text{if } \xi_{jt} < C_f + \sum_{k=1}^{K} \gamma_{1k} VI_{kit} + \gamma_2 r_{it} + \gamma_3 b_{it}, \\ 0 & \text{otherwise,} \end{cases} \tag{1.15}$$

where

$$
\begin{aligned}
C_f =& C_2 \left[G'\left(1+\bar{\xi}\right)\right]^{-1}, \\
\gamma_{1k} =& \left[G'\left(1+\bar{\xi}\right)\right]^{-1} \left[G(\zeta) - G(1+\bar{r})\right] \theta_{1k}, \quad \forall k \in \{1, 2, \ldots, K\}, \\
\gamma_2 =& \left[G'\left(1+\bar{\xi}\right)\right]^{-1} \left\{\left[1 - \Theta\left(\bar{\mathbf{X}}\right)\right] G'(1+\bar{r}) + \left[G(\zeta) - G(1+\bar{r})\right] \theta_2\right\}, \\
\gamma_3 =& \left[G'\left(1+\bar{\xi}\right)\right]^{-1} \left[G(\zeta) - G(1+\bar{r})\right] \theta_3.
\end{aligned}
$$

And let's denote $\xi_{it} \equiv C_f + \sum_{k=1}^{K} \gamma_{1k} VI_{kit} + \gamma_2 r_{it} + \gamma_3 b_{it}$, which depends on i and t only and is independent of j. Thus $\Pr(F_{ijt} = 1) = \Xi(\xi_{it})$. In addition, because $G(\cdot)$ is a concave utility function and $\zeta < 1 + \bar{r}$, it implies $G'(\cdot) \geq 0$ and $G(\zeta) < G(1+\bar{r})$ and hence γ_{1k} and θ_{1k} have opposite signs, i.e., $\operatorname{sgn}(\gamma_{1k}) = -\operatorname{sgn}(\theta_{1k})$ for all $k \in \{1, 2, \ldots, K\}$. □

In the end, loan i is successfully funded if and only if it becomes fully funded. That is at least b_{it} lenders are willing to fund it, i.e.,

$$
F_{it} = \begin{cases} 1 & \text{if } \int_0^{L_t n_{it}} F_{ijt} dj \geq b_{it}, \\ 0 & \text{otherwise.} \end{cases} \tag{1.16}
$$

If more than b_{it} investors would like to invest in loan i at period t, then b_{it} of them are selected randomly to fund the loan.[14]

Because there is a continuum of potential lenders, and the funding decision is symmetric for all potential lenders, combining Equation (1.15) and (1.16) implies that

$$
F_{it} = \begin{cases} 1 & \text{if } L_t n_{it} \Xi(\xi_{it}) \geq b_{it}, \\ 0 & \text{otherwise,} \end{cases} \tag{1.17}
$$

where $\xi_{it} = C_f + \sum_{k=1}^{K} \gamma_{1k} VI_{kit} + \gamma_2 r_{it} + \gamma_3 b_{it}$ and $n_{it} = \frac{1}{N_t} + \pi_{it}$ with $\sum_{i=1}^{N_t} \pi_{it} = 0$ Equation

[14] In practice, P2P market adopts a first come first serve rule. Interested lenders can keep purchasing loan shares until the requested amount is reached and hence the loan becomes successfully funded. However, this is equivalent to the assumption of random selection made in the model conditional on the arrival of potential investors is random.

(1.17) is equivalent to

$$F_{it} = \begin{cases} 1 & \text{if } \ln L_t + \ln n_{it} + \ln \Xi\left(\xi_{it}\right) - \ln b_{it} \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

The first-order approximation (with respect to $\mathbf{VI}_{it}$, r_{it} and b_{it}) of $\ln L_t + \ln n_{it} + \ln \Xi\left(\xi_{it}\right) - \ln b_{it}$ is

$$\begin{aligned} &\ln L_t + \ln n_{it} + \ln \Xi\left(\xi_{it}\right) - \ln b_{it} \\ \approx &\alpha_0 + \sum_{k=1}^{K} \alpha_{1k} VI_{kit} + \alpha_2 r_{it} + \alpha_3 b_{it} + u_{it}, \end{aligned} \tag{1.18}$$

where

$$\begin{aligned} \alpha_0 =& 1 + \ln \Xi\left(\bar{\xi}\right) - \ln \bar{b} - \left[\Xi\left(\bar{\xi}\right)\right]^{-1} \Xi'\left(\bar{\xi}\right)\left(\bar{\xi} - C_f\right), \\ \alpha_{1k} =& \left[\Xi\left(\bar{\xi}\right)\right]^{-1} \Xi'\left(\bar{\xi}\right) \gamma_{1k}, \quad \forall k \in \{1, 2, \ldots, K\}, \\ \alpha_2 =& \left[\Xi\left(\bar{\xi}\right)\right]^{-1} \Xi'\left(\bar{\xi}\right) \gamma_2, \\ \alpha_3 =& \left[\Xi\left(\bar{\xi}\right)\right]^{-1} \Xi'\left(\bar{\xi}\right) \gamma_3 - \bar{b}^{-1}, \\ u_{it} =& \ln L_t + \ln n_{it}, \end{aligned}$$

and $\bar{\xi} = C_f + \sum_{k=1}^{K} \gamma_{1k} \bar{VI} + \gamma_2 \bar{r} + \gamma_3 \bar{b}$.

Thus, the funding equation of loan i can be expressed in a linearized form as

$$F_{it} = \begin{cases} 1 & \text{if } \alpha_0 + \sum_{k=1}^{K} \alpha_{1k} VI_{kit} + \alpha_2 r_{it} + \alpha_3 b_{it} + u_{it} \geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{1.19}$$

where u_{it} increases as either L_t or n_{it} increases. And let's recall that $n_{it} \equiv \frac{1}{N_t} + \pi_{it}$, thus u_{it} also decreases in N_t. Moreover, because $\Xi\left(s\right)$ and $\Xi'\left(s\right)$ are the cumulative distribution function and the probability density function respectively, it implies $\Xi\left(s\right) \in [0, 1]$ and $\Xi'\left(s\right) \geq 0$ and hence α_{1k} and γ_{1k} share the same sign, i.e., $\operatorname{sgn}\left(\alpha_{1k}\right) = \operatorname{sgn}\left(\gamma_{1k}\right)$ for all $k \in \{1, 2, \ldots, K\}$. The derived two linear equations (1.19) and (1.7) can be used directly in the empirical study as funding and default equations.

1.6 Empirical Strategy

1.6.1 Research Question and Empirical Challenges

The key research question is to exam whether P2P lenders form correct expectations of the borrower's default probability conditional on the information available. I focus on the observed borrower information verification in particular, through the comparison between coefficients $\tilde{\beta}_{1k}$ in Equation (1.10) and θ_{1k} in Equation (1.11) for all $k \in \{1, 2, \ldots, K\}$. $\tilde{\beta}_{1k}$ and θ_{1k} can be interpreted respectively as the true and the lender's perceived contribution of verified information k on default probability. For example, suppose that $\tilde{\beta}_{1k} < 0$ and $\theta_{1k} < 0$ for some $k \in \{1, 2, \ldots, K\}$.[15] That means, other things equal, a borrower who has verified information k should have a lower likelihood to default on average compared to another borrower who has not verified the information, and this is perceived by also the potential lenders in the P2P market.

The answer to the research question has strong implications for the efficiency of the P2P market. If equation $\tilde{\beta}_{1k} = \theta_{1k}$ holds for all $k \in \{1, 2, \ldots, K\}$, it implies that the lender's subjective expectation is the same as the objective expectation and hence the investors are likely to make efficient funding decisions. However, if $\tilde{\beta}_{1k} > \theta_{1k}$ (or $\tilde{\beta}_{1k} < \theta_{1k}$) is true for some k, then lenders make inefficient funding judgments by overestimating (or underestimating) the importance of the verification of information k. Overestimation (or underestimation) occurs when lenders' subjective expectations predict a default risk lower (or higher) than that is predicted by the true model, i.e., when lenders overestimate (or underestimate) the safeness of loans.

However, there exist three big empirical challenges to making valid comparisons. The first challenge is the unobserved factors that affect the borrowers' choices of information verification. The observed information verification status is the result of borrowers' optimization problems that involve tradeoffs between the benefits and costs of verification. For example, a borrower who is in a desperate need of funds is more likely to verify more personal information because this can potentially increase the funding probability. On the other hand, it is costlier for some borrowers, such as freelancers, to verify job and income related information because of an unstable working status. Unfortunately, these benefits and costs are unobservable and differ among borrowers. This may result in omitted variable biases when study borrowers' choices of information verification. However, the concern of unobservables is resolved when study the funding

[15]This hypothesis is to be empirically tested in the following section.

decisions from the perspective of lenders because I observe all the borrower characteristics that are observed by lenders and affect lenders' funding decisions. Second, instead of the lender's underlying subjective expectation as displayed in Equation (1.11), as a researcher, I only observe their revealed funding decision. Third, only successfully funded loans have records of repayment or default. I will discuss the solutions to the second and third challenges in detail in the following sections.

1.6.2 Testable Implications

To address the second empirical challenge, it is important to recall that the ratio of α_{1k}/θ_{1k} is a negative constant for all $k \in \{1, 2, \ldots, K\}$. In other words, $\alpha_{1k} > 0$ implies $\theta_{1k} < 0$ and vice versa. Though it is hard to directly and quantitatively compare $\tilde{\beta}_{1k}$ and θ_{1k} , it is possible to compare $\tilde{\beta}_{1k}$ and α_{1k} instead in a qualitative way.

Suppose the lender's subjective expectation of loan default probability is the same as the objective one, then it must imply the following testable properties: if $\tilde{\beta}_{1k} < 0$ for verified information k, then we ought to have $\theta_{1k} < 0$ and hence $\alpha_{1k} > 0$. That is, if the verification of information k implies a significantly lower loan default probability, then *ex ante* the lenders should prefer the borrowers who have verified information k. Alternatively, if $\tilde{\beta}_{1k} = 0$ (or $\tilde{\beta}_{1k} > 0$), then it implies $\alpha_{1k} = 0$ (or $\alpha_{1k} < 0$). Namely, the comparison between the signs of $\tilde{\beta}_{1k}$ and α_{1k} constitutes a feasible test of the necessary conditions for the hypothesis that the P2P lenders are making correct funding decisions.

1.6.3 Sample Selection Issues

The third challenge is a typical example of the well-known problem of sample selection in the empirical literature. Without correcting for selection bias, the estimate of $\tilde{\beta}_{1k}$ is likely to be biased. To highlight the sample selection problem, we can fix all the variables in the loan information set $\mathbf{X}_{it}$ to be constants except for the verification indicator of information κ, $VI_{\kappa it}$.

Then the funding and default equations, Equation (1.7) and (1.19), can be simplified to

$$D_{it} = \tilde{\beta}_0 + \tilde{\beta}_{1\kappa} VI_{\kappa it} + \tilde{D} + e_{it},$$

$$F_{it} = \begin{cases} 1 & \text{if } \alpha_0 + \alpha_{1\kappa} VI_{\kappa it} + \tilde{F} + u_{it} \geq 0 \\ 0 & \text{otherwise} \end{cases},$$

where $\tilde{D} = \sum_{k \neq \kappa}^{K} \tilde{\beta}_{1k} VI_k + \tilde{\beta}_2 r + \tilde{\beta}_3 b$ and $\tilde{F} = \sum_{k \neq \kappa}^{K} \alpha_{1k} VI_k + \alpha_2 r + \alpha_3 b$ are constants.

However, due to the sample selection issue, instead of D_{it}, only D^*_{it} is observed which is defined as

$$D^*_{it} = \begin{cases} D_{it} & \text{if } F_{it} = 1, \\ NA & \text{otherwise}, \end{cases}$$

where NA reflects that the observation is not available. Thus, the regression of D^*_{it} on VI_{it} results in a biased estimate for $\tilde{\beta}_{1\kappa}$ as long as (i) e_{it} and u_{it} are not independent of one another, and (ii) $\alpha_{1\kappa} \neq 0$. This is manifested as follows.

$$\begin{aligned} \tilde{b}_{1\kappa} =& \mathbb{E}\left(D^*_{it} | VI_{\kappa it} = 1\right) - \mathbb{E}\left(D^*_{it} | VI_{\kappa it} = 0\right) \\ =& \mathbb{E}\left(D_{it} | VI_{\kappa it} = 1, u_{it} \geq -\alpha_0 - \alpha_{1\kappa} - \tilde{F}\right) - \mathbb{E}\left(D_{it} | VI_{\kappa it} = 0, u_{it} \geq -\alpha_0 - \tilde{F}\right) \\ =& \tilde{\beta}_{1\kappa} + \mathbb{E}\left(e_{it} | u_{it} \geq -\alpha_0 - \alpha_{1\kappa} - \tilde{F}\right) - \mathbb{E}\left(e_{it} | u_{it} \geq -\alpha_0 - \tilde{F}\right), \end{aligned} \tag{1.20}$$

where $\tilde{b}_{1\kappa}$ denotes the estimate for $\tilde{\beta}_{1\kappa}$ and the bias is the difference $\mathbb{E}\left(e_{it} | u_{it} \geq -\alpha_0 - \alpha_{1\kappa} - \tilde{F}\right) - \mathbb{E}\left(e_{it} | u_{it} \geq -\alpha_0 - \tilde{F}\right)$.

Because both e and u are unobserved in the data, it is hard to conclude whether they are independent or not. Therefore, I only consider the estimation of $\tilde{\beta}_{1\kappa}$ is free of selection bias if and only if $\alpha_{1\kappa} = 0$. In the case of $\alpha_{1\kappa} \neq 0$, I propose two potential solutions to resolve this sample selection issue and present estimates for each in Section 2.5 and 2.6.

Solution I: An Instrument for Funding

Since only funded loans have records of repayment and default, the sample selection bias is essentially the specification error in the conditional expectation

$$\begin{aligned}
&\mathbb{E}\left(D_{it}^{*}|\mathbf{VI}_{it}, r_{it}, b_{it}\right) \\
=&\mathbb{E}\left(D_{it}|\mathbf{VI}_{it}, r_{it}, b_{it}, F_{it}=1\right) \\
=&\tilde{\beta}_0 + \sum_{k=1}^{K} \tilde{\beta}_{1k} VI_{kit} + \tilde{\beta}_2 r_{it} + \tilde{\beta}_3 b_{it} \\
&+ \mathbb{E}\left(e_{it}|\alpha_0 + \sum_{k=1}^{K} \alpha_{1k} VI_{kit} + \alpha_2 r_{it} + \alpha_3 b_{it} + u_{it} \geq 0\right).
\end{aligned}$$

Heckman (1979) proposes a solution to correct the selection bias by assuming the error terms e and u are jointly normal distributed. However, this method also requires finding an exogenous instrument that is correlated with selection but does not directly impact the outcome variable. That is, in the context of this model, the instrument should have a significant effect on the loan funding probability though it is independent of the specific borrower and loan characteristics. Namely, the instrument is variable z_{it} such that when it is included in the default and funding equations

$$D_{it} = \tilde{\beta}_0 + \sum_{k=1}^{K} \tilde{\beta}_{1k} VI_{kit} + \tilde{\beta}_2 r_{it} + \tilde{\beta}_3 b_{it} + \rho_d z_{it} + e_{it},$$

$$F_{it} = \begin{cases} 1 & \text{if } \alpha_0 + \sum_{k=1}^{K} \alpha_{1k} VI_{kit} + \alpha_2 r_{it} + \alpha_3 b_{it} + \rho_f z_{it} + u_{it} \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

the associated coefficient ρ_d is zero while coefficient ρ_f is not.

A compelling candidate for this "instrument" would be the number of contemporaneous loan request posts, N_t. In practice, N_t can be measured as the number of loan listings posted within a small time window. It has already been shown in the theoretic model that the error term, u_{it}, in the funding equation, Equation (1.19), is a function of two factors: the total number of the potential lenders, L_t, and the matching probability, n_{it}. $loan_{it}$ has a higher funding probability when it is matched with more potential lenders (i.e., with higher L_t and/or n_{it}). Though there is no available data on L_t and n_{it}, it is assumed that the average matching probability is negatively correlated with N_t.[16] The underlying intuition is that, other things being the same, loan funding probability decreases when more loans are posted at the same time due to a competition effect.

Solution II: Bounded Treatment Effects

Another approach to dealing with the sample selection issue is to construct bounded treatment effects, defining the lowest and highest values that the true treatment effects could be. Because the selection process is neither random nor exogenous, the post-selection control group (e.g., the loans with $VI_k = 0$) is not the correct counterfactual for the post-selection treatment group (e.g., the loans with $VI_k = 1$). This invalidates the estimation of the exact treatment effects with the entire post-selection sample. Following the method proposed in Lee (2009), I estimate the bounded treatment effect that applies to the "always-funded" borrowers who are funded regardless of their information verification status. The intuition is that "always-funded" borrowers appear in both the post-selection control and treatment groups and can be considered as true counterfactuals for one another as comparison to the "marginal" borrowers who are funded if and only if they have information verified and hence appear in the post-selection treatment group only.

The method can only be used to estimate the bounds of individual coefficients one-by-one, and it requires the following assumptions. For any $\kappa \in \{1, 2, \ldots, K\}$, denote $\mathbf{V}_{it}^{\kappa}$ as the column vector that includes everything in $\mathbf{X}_{it}$ except for the variable $VI_{\kappa it}$ and $\chi_{VI_\kappa=1}$ as the set of all the possible values of $\mathbf{V}_{it}^{\kappa}$ conditional $VI_{\kappa it} = 1$.

Assumption 1.1. *Conditional Independence:* $\{D_{it}(0), F_{it}(0)\} \perp VI_{\kappa it} | \mathbf{V}_{it}^{\kappa} = \mathbf{v} \;\; \forall \mathbf{v} \in \chi_{VI_\kappa=1}$, *where* $D_{it}(1)$ *and* $F_{it}(1)$ *(and* $D_{it}(0)$ *and* $F_{it}(0)$*) denote respectively the potential default and funding probabilities of loan* i *posted at period* t *conditional on the borrower having (and having not) verified information* κ*, i.e.,* $VI_{\kappa it} = 1$ *(and i.e.,* $VI_{\kappa it} = 0$*).*

Assumption 1.2. *Common Support:* $\Pr(VI_{\kappa it} = 1 | \mathbf{V}_{it}^{\kappa} = \mathbf{v}) < 1, \quad \forall \mathbf{v} \in \chi_{VI_\kappa=1}$.

Assumption 1.3. *Joint Independence: The error terms* (e_{it}, u_{it}) *are jointly independent of* $VI_{\kappa it}$ *conditional on* $\mathbf{V}_{it}^{\kappa} = \mathbf{v} \;\; \forall \mathbf{v} \in \chi_{VI_\kappa=1}$.

[16] In the model, L_t, N_t and n_{it} are assumed to be given exogenously. In reality, the variables may be predictable and endogenously determined by the development of the whole P2P market and the overall economy in the long run. For instance, if the market keeps being proved to provide borrowers with easy and low-cost access to well-needed funds and lenders with good-quality investment opportunities, more people will join the market and hence L_t and N_t will increase overtime. However, within a short period of time, e.g., as short as the hourly window studied in the empirical part, L_t, N_t and n_{it} can be considered as exogenous.

Assumption 1.1 and 1.2 are made following Lechner and Melly (2010) to deal with the issue that the choice of borrower information verification is not decided randomly and $VI_{\kappa it}$ is most likely correlated with VI_{kit} for any $k \in \{1, 2, \ldots, K\}$.[17] To isolate the effect of verifying information κ from the effects of verifying the other information, one has to compare the individual loans that have the same lenders-observable loan information set except for the choice of verifying information κ. The assumption of conditional independence implies that if two loans are almost the same except that one has verified information κ while the other has not, then the default and funding probabilities of one without verifying information κ can be considered as the potential counterfactuals for the other. The assumption of common support requires that for a group of loans with information κ verified, there must exist at least another loan without verifying information κ (so that it can be used as the control group) conditional on having all the other observable information being the same.

Provided that Assumptions 1.1 and 1.2 are satisfied, Assumption 1.3 (it is already satisfied given the assumptions made in the model) enables the estimation of the bounded treatment effect. As implied in the derivation of the selection bias, Equation (1.20), the identification of $\tilde{\beta}_{1\kappa}$ would be feasible if I could estimate

$$\mathbb{E}\left(D^*_{it}|VI_{\kappa it} = 1, u_{it} \geq -\alpha_0 - \tilde{F}\right) = \tilde{\beta}_0 + \tilde{\beta}_{1\kappa} + dd + \mathbb{E}\left(e_{it}|u_{it} \geq -\alpha_0 - \tilde{F}\right) \tag{1.21}$$

instead of $\mathbb{E}\left(D^*_{it}|VI_{\kappa it} = 1\right)$. Equation (1.21) is the expected default probability of a subset of funded loans with $V_\kappa = 1$. These loans in the subset would also be funded if they have $V_\kappa = 0$ instead. This is guaranteed by the condition $u_{it} \geq -\alpha_0 - \tilde{F}$, the same condition that determines whether a loan with $V_\kappa = 0$ is funded. If the expected default probability in Equation (1.21) is observable, $\tilde{\beta}_{1\kappa}$ can be readily estimated as the difference

$$\tilde{\beta}_{1\kappa} = \mathbb{E}\left(D^*_{it}|VI_{\kappa it} = 1, u_{it} \geq -\alpha_0 - \tilde{F}\right) - \mathbb{E}\left(D^*_{it}|VI_{\kappa it} = 0\right)$$

Unfortunately, it is not directly observed; however, its value can be bounded.

The set of the funded loans with $VI_{\kappa it} = 1$ can be divided into two subsets: (i) the loans

[17] For example, in reality, a borrower who chooses to verify his job information has a high likelihood of verifying his income information as well.

that are funded only if $VI_{\kappa it} = 1$ and therefore defined as the "marginal" loans, and (ii) the "always-funded" loans. Consequently, the expected default probability of the funded loans with $VI_{\kappa it} = 1$ can be expressed as the weighted average between the expected default probabilities of the two subsets, i.e.,

$$\begin{aligned}\mathbb{E}\left(D_{it}^{*}|VI_{\kappa it}=1\right) =&(1-q)\mathbb{E}\left(D_{it}^{*}|VI_{\kappa it}=1, u_{it}\geq -\alpha_0-\tilde{F}\right)\\ &+q\mathbb{E}\left(D_{it}^{*}|VI_{\kappa it}=1, -\alpha_0-\alpha_{1\kappa}-\tilde{F}\leq u_{it}< -\alpha_0-\tilde{F}\right)\end{aligned}$$

where $q = \Pr\left[-\alpha_0-\alpha_{1\kappa}-\tilde{F}\leq u_{it}< -\alpha_0-\tilde{F}\right]/\Pr\left[u_{it}\geq -\alpha_0-\tilde{F}\right]$, is the proportion of the "marginal" loans among the funded loans with $VI_{\kappa} = 1$.

The bounds of $\mathbb{E}\left(D_{it}^{*}|VI_{\kappa it}=1, u_{it}\geq -\alpha_0-\tilde{F}\right)$ are given accordingly

$$\begin{aligned}\mathbb{E}\left[D_{it}^{*}|VI_{\kappa it}=1, D_{it}\leq D_{VI_{\kappa}=1}\left(1-q\right)\right] \leq \mathbb{E}\left(D_{it}^{*}|VI_{\kappa it}=1, u_{it}\geq -\alpha_0-\tilde{F}\right)&\\ \leq \mathbb{E}\left[D_{it}^{*}|VI_{\kappa it}=1, D_{it}\geq D_{VI_{\kappa}=1}\left(q\right)\right],&\quad (1.22)\end{aligned}$$

where $D_{VI_{\kappa}=1}\left(q\right)$ is the q^{th} percentile of the observed default records conditional on $VI_{\kappa} = 1$. And the coefficient of interest, $\tilde{\beta}_{1\kappa}$, must lie in the interval $\left[\tilde{\beta}_{1\kappa}^{LB}, \tilde{\beta}_{1\kappa}^{UB}\right]$, which is determined by the following two differences

$$\begin{aligned}\tilde{\beta}_{1\kappa}^{LB} &= \mathbb{E}\left[D_{it}^{*}|VI_{\kappa it}=1, D_{it}\leq D_{VI_{\kappa}=1}\left(1-q\right)\right] - \mathbb{E}\left(D_{it}^{*}|VI_{\kappa it}=0\right),\\ \tilde{\beta}_{1\kappa}^{UB} &= \mathbb{E}\left[D_{it}^{*}|VI_{\kappa it}=1, D_{it}\geq D_{VI_{\kappa}=1}\left(q\right)\right] - \mathbb{E}\left(D_{it}^{*}|VI_{\kappa it}=0\right).\end{aligned}$$

Though the "always-funded" loans are known to comprise $1 - q$ fraction of the funded loans with $VI_{\kappa} = 1$, it is impossible to identify the exact ones and their default rates within the treatment group. Therefore, the boundaries of the treatment effect are calculated conditional on two extreme cases: (i) it assumes that the "always-funded" borrowers in the treatment group have lower default rates than the "marginal" borrowers, which produces the lower bound $\tilde{\beta}_{1\kappa}^{LB}$; and (ii) it assumes the "always-funded" borrowers in the treatment group have higher default rates, which derives the higher bound $\tilde{\beta}_{1\kappa}^{UB}$.

1.7 Conclusion

In this chapter, I develop a simple theoretical model of P2P lending. The design of the model helps to illustrate the behavior of lenders and borrowers in the P2P market and to derive the reduced form models for empirical analysis.

In the model, a successfully funded borrower decides whether or not to default, trading off between a random liquidity shock and default cost. Though the liquidity shock and default cost are not directly observable by lenders or anyone other than the borrower himself, their distributions are assumed to, at least partially, correlate with observed borrower and loan characteristics – a simplified representation of the information structure in the real market. *Ex ante*, lenders decide whether or not to fund the loan, comparing between the expected return of the loan and random opportunity costs. The funding decision depends on lenders' subjective expectations of loan default risk, which is formulated based on the observed borrower characteristics and loan information. One can test empirically the efficiency of P2P lenders by comparing the subjective expectation of default with the objective one.

In addition, the issue of sample selection bias in empirical analysis is examined. Because loans are not funded randomly and default records are available for funded loans only, estimation of the objective expectation of default can be biased if sample selection issue is not addressed. Potential solutions include using loan competition as exogenous variation for selection in a Heckman selection model and estimating bounded treatment effects.

Simplifications are made in the model set-up to focus on the analysis of funding efficiency. In particular, the model does not specify borrowers' choices of characteristics to present in loan listings or (repeated) interactions between lenders and borrowers on online platforms. Borrower characteristics and loan information are given exogenously from the lender's perspective in the model, but they are likely to be determined endogenously in reality as borrowers maximize funding probability ex ante.

Figures

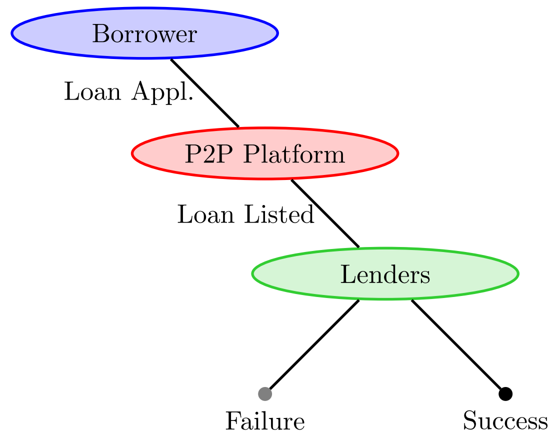


Figure 1.1. A Typical Procedure of P2P Lending

References

Bose, Pinaki. 1998. "Formal-Informal Sector Interaction in Rural Credit Markets." *Journal of Development Economics*, 56: 265–280.

Demyanyk, Yuliya, and Daniel Kolliner. 2014. "Peer-to-Peer Lending is Poised to Grow." Federal Reserve Bank of Cleveland report.

Djankov, Simeon, Caralee Mcliesh, and Andrei Shleifer. 2007. "Private Credit in 129 Countries." *Journal of Financial Economics*, 84(2): 299–329.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig. 2017. "Cultural Proximity and Loan Outcomes." *American Economic Review*, 107(2): 457–492.

Flannery, Mark J. 1986. "Asymmetric Information and Risky Debt Maturity Choice." *The Journal of Finance*, 41(1): 19–37.

Garvey, Kieran, Bryan Zhang, Deborah Ralston, Kong Ying, Rodney Maddock, Hung-Yi Chen, Edward Buckingham, Yianni Katiforis, Luke Deer, Tania Ziegler, Ben Shenglin, Zheng Xinwei, Pawee Jenweeranon, Jingxuan Zhang, Wenwei Li, Rui Hao, and Eva Huang. 2017*a*. "Cultivating Growth: The 2nd Asia Pacific Region Alternative Finance Industry Report." University of Cambridge and Monash University.

Garvey, Kieran, Tania Ziegler, E.J. Reedy, Annie Le, Bryan Zhang, and Randall S. Kroszner and. 2017*b*. "Hitting Stride: 2017 The Americas Alternative Finance Industry Report." University of Cambridge and The University of Chicago.

Garvey, Kieran, Tania Ziegler, Rotem Shneor, Karsten Wenzlaff, Nikos Yerolemou, Rui Hao, and Bryan Zhang. 2017*c*. "Expanding Horizons: The 3rd European Alternative Finance Industry Report." University of Cambridge.

Greenbaum, Stuart I., George Kanatas, and Itzhak Venezia. 1989. "Equilibrium Loan Pricing under the Bank-Client Relationship." *Journal of Banking and Finance*, 13: 221–235.

Haliassos, Michael. 2013. *Financial Innovation: Too Much or Too Little?* MIT Press.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1): 153–161.

Inderst, Roman, and Holger M. Mueller. 2007. "A Lender-based Theory of Collateral." *Journal of Financial Economics*, 84: 826–859.

Jappelli, Tullio, and Marco Pagano. 2002. "Information Sharing, Lending and Defaults: Cross-Country Evidence." *Journal of Banking and Finance*, 26: 2017–2045.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil. 2014. "Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency." *The Review of Economics and Statistics*, 96(1): 1–18.

Kim, Moshe, Jordi Surroca, and Josep A. Tribo. 2014. "Impact of Ethical Behavior on Syndicated Loan Rates." *Journal of Banking and Finance*, 38: 122–144.

Lechner, Michael, and Blaise Melly. 2010. "Partial Identification of Wage Effects of Training Programs." *SSRN Working Paper, 1596715.*

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*, 76: 1071–1102.

Morrison, Alan D. 2005. "Credit Derivatives, Disintermediation, and Investment Decision." *The Journal of Business*, 78(2): 621–648.

Niinimaki, J.-P. 2009. "Does Collateral Fuel Moral Hazzard in Banking?" *Journal of Banking and Finance*, 33: 514–521.

Pevzner, Mikhail, Fei Xie, and Xiangang Xin. 2015. "When Firms Talk, Do Investors Listen? The Role of Trust in Stock Market Reactions to Corporate Earnings Announcements." *Journal of Financial Economics*, 117: 190–223.

PwC. 2015. "Peer Pressure: How peer-to-Peer lending platforms are transforming the consumer lending industry." PwC report.

Schenone, Carola. 2010. "Lending Relationships and Information Rents: Do Banks Exploit Their Information Advantages?" *The Review of Financial Studies*, 23(3): 1149–1199.

Sharpe, Steven A. 1990. "Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships." *The Journal of Finance*, 45(4): 1069–1087.

Srethapramote, Smittipon, Huw van Steenis, Betsy Graseck, Fiona Simpson, Richard Xu, Richard Wiles, Cheryl Pate, Ken Zerbe, Vasundhara Govil, Danyal Hussain, Jeffrey Adelson, Steven Wald, Andrea Ferraz, Denny Galindo, and Michael Cyprys. 2015. "Global Marketplace Lending Disruptive Innovation in Financials." Morgan Stanley Research Global.

Stiglitz, Joseph E., and Andrew Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." *The American Economic Review*, 71(3): 393–410.

Stiglitz, Joseph E., and Andrew Weiss. 1983. "Incentive Effect of Termination: Applications to the Credit and Labor Markets." *The American Economic Review*, 73(5): 912–927.

Stiglitz, Joseph E., and Andrew Weiss. 1992. "Asymmetric Information in Credit Markets and Its Implications for Macro-Economics." *Oxford Economic Papers, New Series*, 44(4): 694–724. Special Issue on Financial Markets, Institutions and Policy.

Chapter 2

Do Lenders Value the Right Characteristics? Evidence from Peer-to-Peer Lending

2.1 Introduction

In this chapter, I investigate two research questions relating to the efficiency of online lending markets using a dataset from Chinese P2P lending platform. First, which borrower traits are crucial to lenders when they make funding decision? Second, do lenders give reasonable expectations for loan default as a function of borrower characteristics? I focus on three major classes of borrower characteristics: education, marital status, and employment information. As proposed in Hollingshead (1975), these are the main factors that measure a person's social status.[1] The variables of interest are the borrower's self-reported education level (e.g., whether the borrower's final degree is a college degree or higher), marital status (e.g., whether or not the borrower is single or married), and employment information (e.g., borrower's monthly income and working industry). Further, I examine the effect of the verification status for each type of borrower

[1]In Hollingshead (1975), the four factors of social status are specifically education, occupation, sex and marital status. The borrower's income and working industry correspond to the factor of occupation. However, the factor of sex is omitted here because a borrower's gender is not immediately observable in the online loan post.

information.[2]

To answer the first question, I regress the P2P loan funding outcomes on the borrower's characteristics to estimate the effects of different characteristics on the loan funding probability. The borrower's characteristics include detailed loan information (e.g., loan amount, interest rate, term and borrowing purpose) and the borrower's personal information (e.g., marital status, education, income etc.). Because the data I use covers all available loan and borrower information, no omitted-variable biases exist in the estimation of impacts of borrower traits on the loan funding probability. The estimated effects can be interpreted as the lender's preferences with respect to borrower characteristics. Lenders are found to favor borrowers who self-report to be married, have education beyond high school, have higher income, and who have verified their employment information. A potential challenge to answering the research question is to what extent loan interest rates assigned by the platform already account for borrower characteristics? That is, to what extent does the loan interest rate proxy for borrower characteristics and other loan information. In practice, I find that (1) interest rates are essentially fixed across large groups of borrowers; and (2) estimates are unchanged when flexibly controlling for interest rates.

To answer the second question, one needs to know whether the preference exhibited at the funding stage optimize subsequent loan performance. Specifically, do the borrowers who are favored by the lenders have lower default probabilities? Utilizing detailed repayment histories of all funded loans, I estimate the association between loan default rate and borrower characteristics. However, these estimates are likely to be biased due to the sample selection issue, since the loan funding decision is not random. To address this issue, I use plausibly exogenous variation in the number of loans posted at the same time as an instrument for the loan funding probability. Specifically, I exploit variation across weeks for loans posted on the same day and during the same hour. This reveals that loans posted at the same time as other loans are significantly less likely to be funded than those facing less competition. The resulting instrumental variables estimates reveal that the loan default rate is significantly lower for borrowers who self-report higher education and who have verified their marital status or education, but are not lower for

[2] There are two variables in the dataset that indicate the verification status of borrower's monthly income and working industry respectively. Because these two variables are extremely highly correlated (the correlation is above 0.95), I combine them into one variable, namely verification of employment information, to avoid problem of multicollinearity. A borrower's employment information is verified if the borrower has verified both monthly income and working industry.

those who have verified employment information. There is no significant association between loan default rate and borrower's self-reported marital status and self-reported income.

Two alternative approaches support these conclusions. First, I restrict attention to borrower characteristics for which there is no evidence of selection in the first stage. Second, I estimate bounds for the treatment effects. I follow the assumptions made by Lee (2009) and Lechner and Melly (2010) to estimate the average bounded effects on treatment to resolve selection issue for the borrower's traits that significantly impact funding probability. In both cases, the results are consistent with those from the instrumental variables design.

Empirical evidence presented in this chapter suggests that the funding decisions made by P2P lenders in the Chinese market are not fully efficient. Lenders show preference for borrowers with verified employment information over those with verified education and marital status despite the predictive power of the later for loan default. As a result, they underestimate the importance of the verification of marital status and education. Borrowers with verified marital status and education are associated with lower default risk, but they are not rewarded with higher funding opportunities *ex ante*. The instrumental variable results indicate that lenders actually overestimate the importance of verifying employment information and self-claims of being married. The findings from both IV and Lee bounds reveal that lenders overestimate having high income on repayment of probability. Regarding self-reported employment industry, borrowers who work in finance, law, real estate, IT, and NGO are underestimated while those who work in construction, energy, and utility are relatively overestimated. With respect to the other characteristics of interest such as the self-claimed education level, the lender's evaluation falls in a reasonable range relative to default. In addition, the coefficient estimates of interaction terms between information verification and self-reported borrower information imply that verification of education does not improve *ex ante* funding opportunity regardless of the borrower's education level, despite being associated with lower default risk for more educated borrowers. On the other hand, verification of employment increases funding probability significantly for borrowers at all income levels, though it is associated with higher default probability *ex post*. Hence lenders underestimate the importance of verifying education but overestimate the importance of verifying employment.

The research contributes to the literature that evaluates lender behavior in the online loan market in several ways.[3] First, unique features of the Chinese market enable me to consider not only the borrower's self-reported information, but also the information verification status in order to estimate the importance of information credibility. Additionally, this study develops a new identification strategy to generate valid estimates of borrower characteristics on loan default. From a broader perspective, this research not only contributes to a better understanding of creditors' expectations in the P2P market, but also sheds light on managerial policies by discussing the potential mechanisms that influence lenders' behavior in the market. My results indicate that interest rates and borrower quality are not highly correlated, and thus interest rates may mislead lenders with respect to default probability. The platform can potentially improve the efficiency of the market by better aligning loan interest rates and default risk.

Much of the literature on P2P lending focuses on borrower characteristics such as gender, social network and cultural proximity, and their effects on the funding outcomes of loan posts (see Pope and Sydnor (2011), Ravina (2012), Agrawal, Catalini and Goldfarb (2015), Freedman and Jin (2008), Lin, Prabhala and Viswanathan (2013) and Kim, Surroca and Tribo (2014) etc.). However, presumably due to data limitations, very little of the literature has examined the role of other borrower social status measures and their influences on funding decision in P2P lending market, and even fewer can compare self-reported with verified borrower information. Moreover, this research is unique in generating unbiased estimates of the effect of borrower characteristics on loan default.

2.2 Literature Review

This work is related to a growing literature on the financial technology sector, including P2P lending, and crowdfunding. Since the first P2P platform, Zopa, was launched in the United Kingdom in 2005, researchers have taken an interest in this new market. Most notably, a small literature has focused on the leading US P2P lending platform Prosper, due to the fact that its data is accessible to the public. Most of the results from the traditional financial sector appear to apply to this newly invented industry. For example, (even unverifiable) information sharing is also critical in determining P2P loan fundings. In the US market, except for credit scores, all information (e.g., borrower's occupation, income, and etc.) displayed in the P2P loan lists

[3] For more literature see Barasinska and Schafer (2014), Duarte, Siegel and Young (2012), Lin and Viswanathan (2016) and Zhang and Liu (2012) etc.

are considered as unverifiable. Nevertheless, Michels (2012) shows that additional information disclosures are associated with lower interest rates and higher bidding activity. Another example is home bias, which is evident in traditional financial markets and for online investment. Burtch, Ghose and Wattal (2014) and Lin and Viswanathan (2016) identify that decreased cultural and geographic distance can facilitate lending using the data from varied crowdfunding platforms. Marom, Robb and Sade (2016) investigate whether crowdfunding reduces the barriers of female entrepreneurs to raise pre-seed capital. They find that men seek significantly higher levels of capital than women who enjoy higher rates of success instead. Barasinska and Schafer (2014), however, find no gender effect on funding probability.

The literature generates several new findings thanks to the unique features of internet financial intermediaries. Firstly, Prosper originally displayed the borrower's appearance in the loan list. Studies such as Ravina (2012) and Duarte, Siegel and Young (2012) utilize this unique feature and find that beauty, race, age, and gender affect lenders' decisions. Consistent with the trust-intensive nature of lending, these studies find that borrowers who appear more trustworthy have higher probabilities of having their loans funded. Social networks built into P2P lending and crowdfunding platforms benefit both creditors and debtors by alleviating information asymmetry through interaction (see Lin, Prabhala and Viswanathan (2013), Agrawal, Catalini and Goldfarb (2015) and Freedman and Jin (2008)). Secondly, in P2P lending and crowdfunding, investors can easily observe the investment behaviors of each other in a timely manner. Instead of passively mimicking their peers (irrational herding), Zhang and Liu (2012) and Liu et al. (2015) find P2P lenders engage in active observational learning (rational herding). Such herding phenomenon is salient in the Chinese P2P lending market as well, but is dominated by passive mimics. Chen and Lin (2014) attribute this difference to cultural and economic factors. Third, several papers have illustrated the role of narratives in P2P funding decisions. Iyer et al. (2015) suggest that screening through soft or nonstandard information allows lenders to have greater accuracy in predicting individuals' likelihood of default. Conversely, Dorfleitner et al. (2016) find that spelling errors, text length, and the mentioning of positive emotion evoking keywords significantly impact funding probabilities but do not predict default probabilities in P2P lending.[4] Finally, Chen and Xu (2015) find that credit insurance has facilitated the P2P market

[4]For more evidence of soft information in the context of online lending, see Sonenshein, Herzenstein and Dholakia (2011) and Herzenstein, Sonenshein and Dholakia (2011).

expansion. Without the well-established financial credit-rating and legal systems in China, online P2P market would be constrained by high credit risk. Consequently, online marketplaces cooperate with offline credit insurers to provide loan guarantees for P2P borrowers.

The study extends the literature by not only estimating the effects of different borrower characteristics on loan determination and performances, but also the effects of information verification. The introduction of information verification is an innovation of the Chinese P2P market to deal with the credit risk. The borrower information considered in this study complements the existing literature by focusing on social status. Furthermore, this research develops a new identification strategy to solve the sample selection issue to have unbiased estimates of borrower characteristics on the loan default.

2.3 Background of Renrendai and Data Description

In this section, I study the funding and default decisions made by P2P lenders and borrowers using the data of Renrendai, one of the best-known and most reputable Chinese P2P lending platforms. The platform was launched in October 2010 and ranked as the 10th largest Chinese P2P lending platform according to total transaction volume in 2014.[5] Up until the second quarter of 2015, the platform has issued over 150,000 loans with a total amount exceeding $1.4 billion. Figure 2.1 displays the remarkable growth in the number and volume of loans consummated via the platform.[6] The number of loans made within a quarter has increased from under 1,000 in the early 2012 to over 25,000 in 2015; the quarterly volume exceed $250 million in the second quarter of 2015.[7]

2.3.1 The Typical Procedure of Borrowing through Renrendai

When applying for P2P loans on Renrendai, borrowers requests loan amount and terms that they would like to borrow and the platform determines the interest rates. In addition, they are required to report certain personal information and describe their reason for borrowing.[8]

[5]Source: Wangdazhijia.

[6]Data source: quarterly and annual reports of Renrendai. All the reports are self issued by Renrendai, and available online at www.renrendai.com/about/about.action?flag=performance. The earliest report is of 2012:Q1 and the latest one is of 2015:Q2. The first observation (of Q4-2011) records the total number and volume of loans that had been made before 2012, which is reported in the quarterly report of 2012:Q1.

[7]We convert Chinese yuan to US dollar with an exchange rate USD/CNY = 0.16/1, that approximates the average exchange rate through the sample period.

[8]The typical procedure of P2P lending is visualized in Figure 2.C.1a in the Appendix.

The borrower information in a request includes demographics (e.g., age, education, marital status), working status (e.g., monthly income, working industry), as well as wealth level (e.g., ownership of assets and properties). Every borrower needs to fill out the online questionnaire as the very first step of filing a loan application.[9] The answers to these questions are collected and disclosed on the loan post webpage as self-reported borrower information.

Additionally, the borrowers can choose to verify some, if not all, of their proclaimed information by submitting supporting documents and materials to the online platform, which has a dedicated department that verifies the authenticity of this documentation.[10] For example, borrowers can verify their education levels by uploading copies of their diplomas in the application. The verification process may take several days to a week in time. The bottom half of the loan post webpage presents the verification status for all verifiable information. Consequently, the potential lenders do not know which information the borrowers have tried to verify, only which information has been successfully verified. After a loan application is completed, the platform will list the loan post online sometime during the following week or two to solicit potential funders. Importantly, due to the delays associated with information verification and loan posting, borrowers have little control over what time of day, or even week their loan is posted. And lenders have no clues regarding the time when upcoming loans will be posted, nor do they have any indications about what the quality of these loans will be.

Figure 2.2 displays a sample webpage of loan posts on the Renrendai platform. On the top panel of the page, it lists the borrowing purpose as "Expand production, new investment" followed by the amount of the loan principal as 30,000 yuan, or approximately \$4,800, the annualized interest rate of 10.00%, and the loan term of 3 months. The self-reported borrower information is displayed in the panel below. The next panel records the information verification status and it shows that the borrower has verified four pieces of information in total (i.e., personal ID, marital status, residence, and mobile phone). A paragraph describing the loan is included at the bottom of the page.

After a loan request is posted online, interested lenders can partially or fully fund the loan by purchasing its loan notes. Each note is worth 50 yuan, approximately 8 dollars, and the lenders

[9]For a majority of the questions, borrowers choose their answers from a drop-down list. However, they can also leave it blank if they choose not to answer the question. As a result, the majority of the variables in the dataset are categorical variables.

[10]See the procedure of borrower information verification in P2P lending visualized in Figure 2.C.1b in the appendix.

can purchase one or multiple notes on a first come, first serve basis. A loan is successfully funded when all of the notes are purchased within 30 days after posting, and the funds will be transferred to the borrower within the next few business days. Otherwise, the loan request fails and the lenders who have purchased the notes are fully refunded. All loans are repaid monthly by the borrowers according to a fixed-rate payment scheme. Payments are made in equal amounts until the principal and interest are paid in full. A borrower can default at any time during the repayment period. If a borrower defaults, the platform will reimburse the creditors a fraction of the unpaid principal in exchange for assuming the right to reclaim the remaining loan repayment from the borrower. In addition, the process of reimbursement can take up to several months, so investors suffer loss, in addition to to other investment opportunities.[11]

The platform makes most of its revenue from service fees charged to borrowers whose loans are successfully funded. The fee rate ranges from 0.55% to 0.88% of the loan principal depending on the borrower's risk level as determined by the platform, which is unknown information to lenders. In general, a borrower is labelled as less risky if he verifies more information. And the platform encourages borrowers to verify as much information as possible by charging a lower rate of service fee so that it can better control loan default risk. Specifically, when a borrower verifies more personal information to the platform, the cost of default increases as the platform has a better chance of finding the person and launching a successful lawsuit against the borrower to reclaim the loan.

2.3.2 Data and Summary Statistics

The data set covers the period from June, 2013, to December, 2014. During this time, the platform was already a top Chinese P2P platform with large and stable daily volumes of borrowers and lenders. The funding rate of loans posted without any information verification is approximately equal to zero.[12] Thus, I restrict the sample to include only loans with at least one

[11]The platform operates an "insurance pool" to reimburse lenders in the case of default. The pool is collected mainly from the management fees charged to the borrowers whose loans are successfully funded. When a scheduled monthly repayment is overdue for more than two weeks, the platform will reach out to the borrower via emails or phone calls. And the loan is considered a default after the payment is overdue for more than three months. According to the platform's policy, lenders of defaulted loans can be reimbursed up to 100% of the unpaid principal on a first-come, first-served basis. However, the actual reimbursements are subject to the availability of the insurance pool and the process usually takes from several months to a year. To refund the lenders, the platform buys out their loan note holdings. This grants the platform the reclamation of the remaining loan payments, which becomes the platform's income if the reclaim succeeds.

[12]The dataset is acquired through web scraping from www.renrendai.com. Namely, it is the snapshot in May 2016 of all the historical loan posts and their associated funding and repayment records. The sample period

piece of verified information. As a result, the dataset comprises 44,765 loan listings of 24,165 distinct borrowers.[13] During the sample period, 8,547 loans (and 7,287 distinct borrowers) were successfully funded, resulting an overall funding rate of 19.1%. Of the funded loans, 18.5% were funded by a single lender, while the rest were funded by multiple lenders. As a result, a funded loan has on average 16 investors. A little more than 10% of the funded loans have records of default within the repayment period.

Table 2.1 shows the summary statistics of the loan contract information, the information verification indicators, and self-reported information for all of the funded and unfunded loans in the sample.[14] For the loan contract information, the interest rates are in annualized percentage rates, the principals are converted to dollars using the exchange rate 1 USD = 6.25 CNY, and the loan terms are in months. The average interest rate in the sample period is 13.71%. As a comparison, during the same period, the prime lending rate of Chinese bank loans varied from 5.60% to 6.55%, while the mortgage rate ranged from 4.86% to 6.94%. For credit card borrowers, the annual interest rate averaged about 15%. Obviously, the borrowing cost is higher from Renrendai than through traditional financial intermediaries. However, the majority of the P2P borrowers are individuals and small and medium-sized enterprises (SMEs) that are likely to have difficulty obtaining funds from the traditional channels due to their lack of collateral, credit, and established relationships with the banks. For lenders, the maximum rate of a five-year certificate of deposit is 4.25%. This helps to explain why investors have an incentive to lend on the P2P market even when loan repayments are not fully guaranteed.

The comparison between the funded and unfunded loans reveals significant differences.[15] For funded loans, the average interest rate is lower by over 120 basis points, or roughly 10 percent. The average loan size is only 27.64% of unfunded requests, and the average term is shorter by 5 months. A much larger fraction of funded borrowers have verified their information.

starts in June 2013 when the platform had already developed into a top Chinese P2P platform with large and stable daily volumes of borrowers and lenders, and it ends in December 2014 so that it has long-enough time for the loan repayments to be observed. During this period, 246,392 loans are posted on Renrendai in total and 10,064 are successfully funded. 62,398 loans are posted without any information verification, and only 3 of them are successfully funded. Apparently, these loans attract little investment interest of the lenders because the borrower information is very likely to be corrupted. To avoid the results being driven by this type of loans, I restrict the sample to exclude the loans with no information verification. Because the dataset is the snapshot of all the information available in May 2016, 139,229 loan posts contain undetermined information verification status such as “being currently under validation”. And these observations are further excluded.

[13]Each loan list on Renrendai platform is labelled by a distinctive loan ID. While the borrowers and lenders are identified by online user IDs.

[14]The self-reported working industry is not included in Table 2.1 because the constrain of the table length and no systematic patterns are found.

[15]The $p-$values indicate the differences are all significant at the 99% confidence level.

The differences are largest for information about employment information, which about 80% of funded borrowers have verified vs. less than 20% of unfunded borrowers. In terms of the self-reported information, funded borrowers are more likely to be married, have monthly income above median ($800 or 5,000 yuan), and have education above the high school level. Since the borrowers do not choose randomly among the information to verify, their choices of information verification can potentially be correlated. However, this is not true according to the table of correlation as shown in Table 2.C.7, where most verification indicators have small correlation with one another.

2.4 Effect of Information Verification on Funding and Default

2.4.1 Empirical Specification

I start the empirical analysis with reduced form regressions as follows. They are derived from Equation (1.19) and Equation (1.7) in the theoretical model to include the same self-reported borrower information as variables of interest as well.

$$F_{it} = \alpha_0 + \sum_{k=1}^{3} \alpha_{1k} VI_{kit} + \sum_{k=1}^{3} \sum_{j=1}^{J(k)} \alpha_{2k}^{j} SR_{kit}(j) + \boldsymbol{\alpha}_3 \boldsymbol{Control}_{it} + u_{it} \tag{2.1}$$

$$D_{it} = \tilde{\beta}_0 + \sum_{k=1}^{3} \tilde{\beta}_{1k} VI_{kit} + \sum_{k=1}^{3} \sum_{j=1}^{J(k)} \tilde{\beta}_{2k}^{j} SR_{kit}(j) + \tilde{\boldsymbol{\beta}}_3 \boldsymbol{Control}_{it} + v_{it} \tag{2.2}$$

where the subscript it indicates loan i posted in time t. F_{it} is a dummy that equals 1 if the loan is funded, and 0 otherwise, while $D_{it} \in [0, 1]$ is a continuous variable that represents the fraction of the $loan_{it}$ that is defaulted. For instance, $D_{it} = 0.5$ in the case of a loan that matures in 12 months whose borrower repays for the first 6 months but defaults on the rest of the payment. In the empirical study, I focus on three categories of borrower information in particular. They are (i) education level, (ii) marital status, and (iii) employment information. Much of the existing literature on P2P lending explore the data from Prosper. Compared to their counterparts on Renrendai, Prosper loan listings contain a relatively smaller set of borrower characteristics. Information regarding a borrower's social status such as marital status and education level is available on Renrendai but not on Prosper. $VI_{kit} = 1$ if borrower information k of $loan_{it}$ is verified, and 0 otherwise. In addition to the information verification status, I am also interested in the corresponding self-reported information, e.g., the final degree that a borrower reports on his education level. Because all self-reported information are categorical variables, I use dummies to represent them. Specifically, I denote $J(k)$ as the set of the different values that the self-reported information k takes, and $SR_{kit}(j) = 1$ if the borrower reports value j, and 0 otherwise. $\boldsymbol{Control}$ denotes a vector of control variables that includes all the other information verification indicators and self-claimed borrower information. The loan contract information includes the loan principal (in dollars), interest rate (in annualized percentage rates), and loan term (in months). Meanwhile, I include monthly fixed effects to control for development in the P2P market and aggregate economic conditions.

Self-reported borrower information take the form of categorical variables, some of which can take many different values. For example, a borrower can report monthly income in 8 different levels (from NA to over $8,000) and there are 20 different working industries. Thus, to simplify the reporting of results, I reclassify the reported information according to the definitions displayed in Table 2.2. In particular, a borrower's marital status is now classified as being either married or unmarried (which includes single, divorced and widowed). The education level of a borrower is divided into three categories: high school and below, associate's degree, and bachelor's degree or higher. Working industries are redefined by the 2-digit North American Industry Classification (NAICS) into 9 different groups. Monthly income is regrouped into two different levels: above and below $800, which is the median income level in the sample.[16] According to the borrower's self-reported city of residence, I also add the city's economic geography and GDP per capita to control for overall economic development.[17]

Besides the hard information discussed above, the self-reported borrower information also includes the soft information contained in the loan description and borrowing purposes. The borrowing purposes of the Renrendai loans can be divided into 10 different categories, which

[16] The results are consistent when monthly income is divided instead into three discrete levels: less than $800, between $800 and $1600, and more than $1600.

[17] Based on the National Bureau of Statistics of China, the country is divided into four economic geographic regions. They are the east coast of China, central China, northeast China and western China. Each region comprises neighboring provinces and share similar economic and industrial structures. The later three regions are defined in and benefit from the different state economic policies adopted by the Chinese central government. Central China is in the Rise of Central China Plan announced in March 2004. Northeast China is in the Northeast Area Revitalization Plan started in late 2003. Western China is in the China Western Development began in 2000. The data of city level GDP per capita is of 2013 and taken from China City Statistical Yearbook 2014.

include different kinds of expenditures, purchases of property and assets, and liquidity demands by businesses. Different purposes may generate different impacts on the funding probability and hence are added into the control variables. A rich literature in financial accounting such as Antweiler and Frank (2004), Engelberg (2008), Li (2008), Tetlock (2007), and Tetlock, Saar-Tsecuansky and Macskassy (2008) use textual analysis to measure the effective sentiments in the corporate statements, which has proven to be significantly correlated with other financial variables. Applying the idea of textual analysis in our context, I compute three different variables to capture the soft information contained in the loan description. The first variable is the log of the number of words in the loan description, which measures the quantity of information. The second variable is the number of top 100 most frequent words used in the description which quantifies the effective information content.[18] The last variable is the number of positive words from the top 100 most frequent words to gauge the sentiment of the description.[19] Among these 100 words, 12 are considered positive words according to the Xin Hua Dictionary and the Xin Hua 08 Chinese-English Financial Dictionary.[20] The connotation of a word is defined in Chinese and may change in the translation from Chinese to English.

2.4.2 Reduced Form Results

able 2.3 and 2.4 display the reduced form results of the funding and default equations estimated using ordinary least squares. The standard errors are clustered by borrower's quality defined by the combination of borrower's choices of information verification and self-claimed responses. In both tables, column (1) shows the estimation results without controlling for the loan interest rate, while in columns (2) to (4), the loan interest rate is included in three different forms: (i) as a continuous variable, (ii) as a discrete factor variable, and (iii) as a discrete factor variable interacted with monthly fix effects. The coefficients on the regressors of interest are very similar in terms of magnitude and statistical significance across the four specifications. In other words, the coefficients are not significantly attenuated by the inclusion of loan interest rate set by the platform. The results are also robust to the inclusion of control variables, including the full set of the verification indicators, self-claimed responses, and the interaction terms between these

[18] I count the frequency of every word used in loan descriptions for the whole sample. The top 100 most frequent words are those with the highest frequencies.

[19] These words are translated into English and listed Appendix 2.A.1.

[20] A detailed list of these positive words is available at Appendix 2.A.1.

two.[21] This implies that the loan interest rate is not set in such a way that it offsets the effects of borrower characteristics. In other words, the interest rates really do not vary much across lenders. In effect, no significant difference is found between the average loan interest rates across borrowers who have verified different types or amounts of information.[22]

The coefficient estimates in Table 2.3 are likely to be an unbiased measure of what characteristics are valued by lenders because all borrower and loan characteristics available to the lender are included in the regression model. The results imply that the P2P lenders consider employment borrower characteristics as the most important statistics when making funding decisions, but ignore other borrower characteristics even though they may contain valuable information regarding default risk. Lenders may prefer borrowers who have higher monthly income and verified employment information because they think they have better income flows and stable employment and can be considered as less risky. However, borrowers are not treated differently in response to their self-reported industry as the industry doesn't necessarily determine one's income level. In terms of education and marital status, it appears that lenders favor borrowers who are married and have higher degree levels, but they do not distinguish between verified and unverified responses. This can be justified if lenders believe borrowers are less likely to make false claims about their marital status and education levels compared to their income levels.

The interpretation of the coefficients in Table 2.4 is conditional on being funded. As proved in Section 1.6.3, the estimate for $\tilde{\beta}_\kappa$ in the default equation is likely to be biased as long as (i) u_{it} and v_{it} are not independent of one another, and (ii) $\alpha_\kappa \neq 0$.[23] Because it is hard to tell the exact components included in the error terms, it becomes impossible to infer whether the two error terms are independent. Instead, it is much easier to test whether selection occurs conditional on information κ. If $\alpha_\kappa = 0$ is found to be true in the data, then information κ has no significant impact on the loan funding probability and does not induce the selection problem. Based on the reduced form results of the funding equation in Table 2.3, the coefficients on the verification of marital status, the verification of education, and the self-claimed working industry are not statistically significant at 90% confidence level. Hence, these variables aren't considered as the causes of sample selection issue. In this case, the coefficient $\tilde{\beta}_\kappa$ in the default equation can be

[21] See the detailed estimation results in Table 2.C.3, 2.C.4, 2.C.5, and 2.C.6 in Appendix 2.C.

[22] See Table 2.C.1 and 2.C.2 in Appendix for the average loan interest rates and associated standard deviations in the sample for borrowers who have verified different types or amounts of information.

[23] α_κ can be considered as either $\alpha_{1\kappa}$ or $\alpha^j_{2\kappa}$ and $\tilde{\beta}_\kappa$ as either $\tilde{\beta}_{1\kappa}$ or $\tilde{\beta}^j_{2\kappa}$ for all $j \in J(\kappa)$ and κ in Equation (2.1) and (2.2).

estimated using the least squares method. The result in Table 2.4 suggests that the borrowers who have verified marital status and education are associated with lower default probability. Relative to those who report working in the manufacturing industry, the borrowers who report working in the construction, energy and utility industries are likely to have higher default rates, while the borrowers who report working in finance, law, real estate, IT, NGO and others are likely to have lower default rates.

As shown in Section 1.6.2, the hypothesis that the lenders make correct funding decisions can be tested through a comparison of the coefficients α_κ with $\tilde{\beta}_\kappa$. The potential comparison outcomes are listed in the first row of Table 2.5 in the case of $\alpha_\kappa = 0$. If $\tilde{\beta}_\kappa > 0$, the lenders overestimate the quality of borrowers with characteristics κ, because the borrowers with characteristics κ have a higher funding probability but their default rates are not lower. Alternatively, if $\tilde{\beta}_\kappa < 0$, then the lenders underestimate the quality of borrowers with characteristics κ. Finally, if $\tilde{\beta}_\kappa = 0$, the lenders are likely to make funding decisions that are consistent with borrower default rates. The comparison between the coefficient estimates in Table 2.3 and 2.4 suggests that lenders underestimate the importance of the borrower's verification of marital status and education. And, relative to borrowers working in manufacturing, lenders overestimate those working in construction, energy, and utilities, while underestimating those working in finance, law, real estate, IT, NGO, and others.[24] For borrower characteristics with statistically significant coefficients in the funding equation, I will apply the instrumental variables and bounding methods to deal with the selection issue.

2.5 Identification Strategy and Instrumental Variable Results

As illustrated in Section 1.6.3, exogenous variation in funding that satisfies the inclusion and exclusion restrictions can be used as an instrumental variable to address the sample selection issue. A candidate is competition from other loans listed at the same time. That is, loans that face less competition are more likely to get funded. This assumption derives from the real experience of investing at Renrendai. On the webpage of browse listings, lenders see a list of the 10 most recent loan posts with only four pieces of information displayed, including the loan

[24]Table 2.C.8 and 2.C.9 in Appendix 2.C exhibit the least squares estimates for the control variables.

amount, interest rate, term and borrowing purpose. If they are interested in any of the 10 loans, they have to click the link to the specific loan post page to view the detailed borrower and loan characteristics before they can make investments. If lenders want to see more loans, they have to click the links to the next pages. Therefore, when more loans are posted around the same time, it becomes harder for any of the loan posts to be viewed and hence invested by potential lenders. The inclusion restriction, that the number of loans and getting funded are correlated, is easily tested.

As exogenous variation, I note that the number of loans released online can vary remarkably within a week and even within a day. As an example, Figure 2.3 displays the number of loans posted each hour of the day for every day in June 2013. The number of loans posted online possesses some predictable patterns. Comparing the number of loans posted at different hours within a day, there are significantly more loans posted during the business hours from 9 am to 7 pm. And relative to weekends, weekdays see slightly more loan posts in general. These features persist during the entire sample period. Therefore, the exogenous variation to be used to deal with the selection issue is the unpredicted variation in the number of loans released during the hourly windows after controlling for the fixed effects of the month, the day of a week and the time within a day. For instance, suppose on average 8 loans are posted between 9 to 10 am on the Mondays in a given month, while 10 loans are posted during the time period on the first Monday, then the exogenous variation is the unpredicted difference of 2 loans.

On the other hand, according to the policy of Renrendai, a loan post is listed online sometime within a week or two after the loan application is completed. Consequently, borrowers have no control over the exact time when their loan requests are posted, and lenders have no information regarding the time when and the quality of new loans to be posted. So it is plausible to assume that the total number of loan posts at a given time is independent of the quality of each loan post. The exclusion restriction requires the number of loan posts has no direct effect on the default probability of individual loans, which is also supported by the data as shown in Section 2.5.

I choose to focus on hourly windows because almost all funded loans are funded within an hour.[25] When loans are posted on Renrendai, they are exhibited on a webpage showing 10 loans

[25]The time is calculated as the difference between the first and last lender bids. For the 8,547 funded loans in the data sample, 99% of them are funded in less than 35 minutes and none of them takes more than 7 hours.

in chronological order from the youngest to the oldest. After logging into Renrendai, lenders see the listing webpage of the 10 most recent loan posts. If they want to view more loans, they have to click the links to the next pages. On average, there are 20 loans posted every hour during the sample period. That is, the typical loan stays on the first page for about 30 minutes. This, echoed by the empirical finding that 99% of funded loans are funded within 35 minutes, implies the "first page" advantage. Intuitively, it is easier for loan lists to solicit more lenders to invest if they can stay on the first page longer. Therefore, if a loan is posted in the hour when relatively more loans are posted contemporaneously, its time on the first page is shorter and its chance to be successfully funded is lower.

The reason to control for monthly fixed effect is to eliminate the anticipated growth trend of the marketplace as well as changes in the overall economy. The sample period witnessed fast development of the Chinese P2P market. The number of loan posts, the volume of facilitated loans, and the population of online borrowers and lenders has been increasing dramatically month by month on leading platforms including Renrendai. The fixed effects for the day of week and the time of day account for weekdays versus weekends and business hours versus non-business hours.

For the aforementioned variation to be a credible instrument, it needs to satisfy the inclusion and exclusion restrictions. Namely, the unpredicted variation has to be significantly correlated with the loan funding probability while independent of the borrower quality. Intuitively, when more loans are posted online and the number of lenders is held fixed, each loan has a lower probability of being funded due to the competition effect. So, the inclusion restriction should be satisfied. And according to the P2P platform policy, the loan lists are posted online at some time after the borrowers submit the loan requests to the platform that is not a function of when the loan was submitted. Therefore, it seems likely that the quality of a loan and its borrower are independent of when the loan is posted and the amount of competition it faces.

As evidence that the inclusion restriction is satisfied, I conduct a regression of Equation (2.3) using least squares,

$$F_{it} = \beta_0 + \beta_1 N_t + \beta_2 Control_{it} + FE_t + \epsilon_{it}, \tag{2.3}$$

where N_t is the total number of loans posted at time t and FE_t includes fixed effects for the month, the day of a week, and the hour in a day. The control variables here include all the explanatory variables in the regression of Equation (2.1). The results are presented in Table 2.6. The effect of the number of loan posts is significantly negative: the loan funding rate decreases by nearly 2% for each additional 10 loans that are posted together at the same time. Thus the result indicates that the inclusion restriction is satisfied. Next, I regress the number of loan posts on the borrower characteristics as in Equation (2.4) to test whether the exclusion restriction is satisfied.

$$N_t = \beta_0 + \beta_1 Control_{it} + FE_t + \epsilon_{it} \tag{2.4}$$

The least squares estimates for the key variables are reported in Table 2.7. The results suggest that there is no significant association between borrower characteristics and the number of loan posts, and hence support the argument that the number of loan posts is valid for being an instrument.

With the number of loan posts as the exogenous variation, Equation (2.1) and (2.2) now constitute the Heckman selection model.[26] The model is estimated using the Maximum Likelihood Estimation (MLE) and the results are reported in Table 2.8. First of all, let us focus on the verification of employment information in addition to the self-reported marital status, education level and monthly income, because these variables are found to have significant impacts on the funding probability by the least squares estimation as shown in Table 2.3. The MLE coefficient estimates imply that the P2P lenders prefer borrowers who have verified their employment information and who are married and have better degrees with higher monthly incomes. The MLE estimation also reports the coefficients for those non-selective borrower characteristics discussed in the previous section. Qualitatively, the findings here are almost the same as those found by the least squares estimation in the the previous section.

The MLE estimates should be unbiased despite the selection issue. This allows us to test whether the lenders make reasonable funding decisions by comparing α_κ and $\tilde{\beta}_\kappa$. The related principles are listed in the second and third rows of Table 2.5 in the case of $\alpha_\kappa \neq 0$. Suppose $\alpha_\kappa < 0$, then the lenders underestimate the borrowers with trait κ if $\tilde{\beta}_\kappa \leq 0$, otherwise they make reasonable funding decisions by rewarding the borrowers with better funding opportunities for they have lower default risks. Instead, suppose $\alpha_\kappa > 0$, then the lenders overestimate the borrowers with characteristics κ if $\tilde{\beta}_\kappa \geq 0$, otherwise their investment decisions are likely

[26]The fixed effects of the day in a week and the hour in a day are added into the control variables so that only the unexpected variation in the number of loan posts are used effectively as the instrument variable.

to be reasonable. Given the estimation results in Table 2.8, one can assert that the lenders overestimate the importance of the borrower's verification of employment information and their self-reported claim of being married and having a monthly income over $800. On the other hand, the lenders make rather reasonable funding decisions with respect to the self-reported education level.

$$F_{it} = \alpha_0 + \sum_{k=1}^{3} \alpha_{1k} VI_{kit} + \sum_{k=1}^{3} \sum_{j=1}^{J(k)} \alpha_{2k}^{j} SR_{kit}\left(j\right) + \sum_{k=1}^{3} \sum_{j=1}^{J(k)} \alpha_{3k}^{j} VI_{kit} \times SR_{kit}\left(j\right) \\ + \boldsymbol{\alpha}_4 \boldsymbol{Control}_{it} + u_{it} \quad (2.5)$$

$$D_{it} = \tilde{\beta}_0 + \sum_{k=1}^{3} \tilde{\beta}_{1k} VI_{kit} + \sum_{k=1}^{3} \sum_{j=1}^{J(k)} \tilde{\beta}_{2k}^{j} SR_{kit}\left(j\right) + \sum_{k=1}^{3} \sum_{j=1}^{J(k)} \tilde{\beta}_{3k}^{j} VI_{kit} \times SR_{kit}\left(j\right) \\ + \tilde{\boldsymbol{\beta}}_4 \boldsymbol{Control}_{it} + v_{it} \quad (2.6)$$

Furthermore, I include interaction terms between information verification and self-reported borrower information into the funding and default equations, as shown in Equation (2.5) and Equation (2.6), to examine whether the effects of information verification on loan funding probability and default rate differ across borrowers with different self-reported information. Table 2.9 lists the estimation results of the funding equation, and the results of the default equation for the selection model using maximum likelihood. I find no effects of verification of marital status on loan funding probability or default rate, which is true regardless of whether a borrower self reports to be married or not. On the other hand, relative to those who self claim to have a high school diploma or below, borrowers who claim to have at least an associates degree (or bachelor's degree) are associated with significantly lower default risk if they have verified their education, though these borrowers are not rewarded with better funding opportunities. Verification of employment information significantly improves borrowers' success rate of getting funded, and this impact is even bigger for borrowers who self claim to have monthly income lower than the median ($800), but the default rate is not lower for these borrowers. In sum, lenders underestimate the effect of verification of education for borrowers with higher self-claimed education levels and overestimate the effect of verification of employment information for borrowers with higher self-claimed monthly income.

2.6 Bounds on the Average Treatment Effects

An alternative to the instrumental variables strategy is to establish lower and upper bounds, as proposed by Lee (2009). The challenge in this research is that the variables of interest are not randomly assigned as in Lee (2009), which studies the wage effects of the Job Corps, a US national job training program. Lee takes advantage of the randomized evaluation of the program in the mid-1990's that randomly divided eligible applicants for the program into the "control" and "treatment" groups so as to assess the wage difference afterwards. However, the differences in the borrowers' characteristics (i.e., the choices of information verification and self-claims of borrower information) are not assigned randomly in my setting, though both have selection; not all of the people studied in Lee's data choose to work, whereas not all of the loans in this dataset are eventually funded. As a compromise, instead of measuring the bounded average treatment effect (ATE) as in Lee (2009) , I estimate the bounded average treatment effects on treated (ATET) for the variables of interest based on Assumption 1.1. For example, let the verification of information κ, VI_κ, be the variable of interest. To satisfy Assumption 1.1, all the loan observations are categorized into quality groups defined by all the borrower characteristics other than VI_κ. Within each quality group, the control observations are those with $VI_\kappa = 0$, while the treatment are those with $VI_\kappa = 1$. Some groups are dropped if they consist of 100 percent control or treatment observations. In most cases, there are far more control than treatment observations within a given quality group. And for the whole sample, the control observations concentrate in some quality groups, while the treatment concentrate in others. To make sure that the distribution of the control observations across all the groups is the same as that of the treatment, the control observations are selected randomly so that the number of control in-use equals the number of treatment in that group.[27] In this way, the control and treatment in-use have the same distributions of borrower and loan characteristics except for the values of VI_κ, and the observed difference in the outcomes can be considered as the result of VI_κ changing from 0 to 1.

Using the selected control and treatment observations, I calculate the bounds on the ATET of

[27] In the cases that there are more treatment than control observations within a group, the treatment observations are selected randomly instead.

verifying information κ on the default rate. As illustrated in Section 1.6.3, the set of the funded loans with $VI_\kappa = 1$ can be divided into two subsets: (i) the "marginal" loans that would not be funded if $VI_\kappa = 0$ and (ii) the "always-funded" loans that are always funded regardless the value of VI_κ. On the other hand, the funded loans with $VI_\kappa = 0$ consist of the "always-funded" loans only. Therefore, the funded control observations can be considered as counterfactual for the "always-funded" treatment.[28] To estimate the ATET for VI_κ, one subtracts the average default rate of the funded loans with $VI_\kappa = 0$ from that of the "always-funded" loans with $VI_\kappa = 1$.

Though it is impossible to identify the "always-funded" from the funded treatment, it is possible to calculate the bounds of the ATET conditional on the two extreme possibilities. Let the proportion of the "marginal" loans among the funded loans with $VI_{\kappa it} = 1$ be denoted by

$$q = \frac{\Pr(F_{it}|VI_{\kappa it} = 1) - \Pr(F_{it}|VI_{\kappa it} = 0)}{\Pr(F_{it}|VI_{\kappa it} = 0)},$$

and the share of the "always-funded" is $1 - q$. In one extreme, all "marginal" loans have lower default rates than the "always-funded" loans, so the average default rate of the "always-funded" loans is the average default rate above the q^{th} quantile of the funded treatment (the upper default rate mean). In the other extreme, all "marginal" loans have higher default rates than the "always-funded" loans, so the average default rate of the "always-funded" loans is the average default rate below the $(1-q)^{th}$ quantile of the funded treatment (the lower default rate mean). Subsequently, the upper and lower bounds of the ATET are calculated as the differences between the upper and lower means and the average default rate of the funded control.[29] This estimation method is used to calculate the ATET bounds of self-reported borrower information, $SR_\kappa(j)$, on the default probability.

When α_κ in the funding equation is found to be statistically different from zero, the selection issue exists and does not allow the direct estimation of $\tilde{\beta}_\kappa$ in the default equation. Instead, $\tilde{\beta}_\kappa^{LB}$ and $\tilde{\beta}_\kappa^{UB}$, the lower and upper boundaries, are calculated and provide an alternation to the IV approach presented in Section 2.5. And, as an alternative method to test whether the lenders make reasonable funding decisions, I compare the coefficient α_κ with the interval $\left[\tilde{\beta}_\kappa^{LB}, \tilde{\beta}_\kappa^{UB}\right]$. Table 2.10 lists the possible outcomes of the comparison. Suppose $\alpha_\kappa > 0$. If the lower boundary $\tilde{\beta}_\kappa^{LB} \geq 0$, then the true effect $\tilde{\beta}_\kappa$ must be positive as well. This implies that a borrower with

[28]The opposite statement may not be true under Assumption 1.1.

[29]The calculation of the corresponding standard errors and confidence intervals can be found in Appendix 2.B.

characteristics κ is more likely to get funded despite having a higher likelihood of default. In other words, lenders overestimate the quality of borrowers with characteristics κ. On the other hand, if $\tilde{\beta}_{\kappa}^{UB} < 0$, then the true effect $\tilde{\beta}_{\kappa}$ is also negative and it is likely that the lenders make reasonable funding decisions by rewarding the borrowers who have characteristics κ with a higher funding probability for their lower default risk. In the opposite situation when $\alpha_{\kappa} < 0$, if $\tilde{\beta}_{\kappa}^{LB} > 0$, then the true effect $\tilde{\beta}_{\kappa}$ is positive. It indicates that the lenders are making accurate funding decisions. Instead, if $\tilde{\beta}_{\kappa}^{UB} \leq 0$, then the true effect $\tilde{\beta}_{\kappa}$ is also negative and it is likely that the lenders underestimate borrowers with characteristic κ.

Furthermore, if the range of the bounded effects is not significantly different from zero, it becomes statistically inconclusive to tell whether the lenders' expectations and their funding decisions are reasonable. Nonetheless, I can still make inference on the inconclusive results as long as the boundaries are relatively narrow. The least squares results in Table 2.3 imply that the higher-quality borrowers (those who verify their job information and have higher education levels and higher income) have a higher funding probability. Intuitively, compared to the "marginal" borrowers, the "always-funded" borrowers are more likely to have a lower default rate. Therefore, one may argue that the default rates should be higher for the "marginal" loans than the "always-funded" loans, so the true ATET should lie closer to the lower bound.

The bounds of the ATET are displayed in Table 2.11 for the variables of interest that have significant impacts on the funding probability. Given the results, it appears that lenders make correct funding decisions with respect to self-reported marital status and education levels, as the upper bounds are less than zero. On the other hand, lenders may overestimate borrowers who claim to have monthly income over \$800, as the upper and lower boundaries are small in magnitude and concentrated around zero. With respect to the verification of employment information, my conjecture is that the lenders make reasonable investment decisions, though the interval includes zero. This is because the upper bound is very close to zero, while the lower bound is far below zero. As a comparison to IV results, these bounds are just a way to double check if the IV estimates are reasonable.

As the robustness check, I also estimate the bounds on the ATE using the full sample as if all the treatments (the verification and self-report of borrower information) were randomly assigned. The results are presented in Table 2.12. The main purpose of this exercise is to test whether my method brings narrower bound of the effects. In general, the bound estimated using the two methods are similar to each other, except for the verification of employment information for which my method generates a narrower boundary.

2.7 Discussion

The results given by the different methods discussed above suggest that the lenders indeed over- or underestimate borrowers with certain characteristics such as verified marital status, education and employment information, and self-reported income and marital status. This results in inefficient funding decisions in the P2P lending market. This section discusses potential mechanisms and intuition behind the phenomenon. I propose several hypotheses to be tested in this section.

2.7.1 Gradual Learning by Lenders

One hypothesis for the underlying phenomenon is that the lenders lack sufficient information and experience when making investment decisions. Consequently, they need time to learn from their mistakes and those of their peers in the market. Such learning may take place gradually, and hence it may take a long time before the lenders are able to make optimal funding decisions. If this hypothesis is true, we are likely to observe evidence that lenders' funding choices improve over time.

To test the hypothesis, I split the dataset into two subsamples with each subsample having a time span of 9 months. The first subsample goes from June 2013 to March 2014, and the second one from April to December 2014. Then the funding equation is reestimated separately for the two subsamples, and the results are included in the second and third columns in Table 2.13. The comparison between the coefficients for the two subsamples finds no significant changes in the impacts on the funding probability for the verification of marital status, education and employment information, which no evidence is found to support the proposed hypothesis of learning by the lenders. However, the finings do not necessarily rule out the possibility that the P2P lenders are learning to improve their funding decisions. The learning process may not be well observed in the data because of the short sample period. Besides, the improvement could be slow due to the cost of learning as the Chinese P2P market is growing rapidly, with hundreds of new platforms launched each year and thousands of new borrowers and lenders joining the market every day.

2.7.2 Misleading Loan Interest Rates

Theoretically, if the loan interest rate fully incorporates the potential default risk, then lenders can ignore borrower characteristics. However, if the interest rate does not account for loan risk, it could potentially mislead the funding decisions. The following Equation (2.7) examines the relationship between the borrower's characteristics and the loan interest rate.

$$r_{it} = \beta_0 + \sum_{k=1}^{4} \beta_{1k} VI_{kit} + \boldsymbol{\beta_2 Control_{it}} + \epsilon_{it} \tag{2.7}$$

The results are reported in the last column in Table 2.13. Though focusing on the three primary information verification indicators, the regression of Equation (2.7) includes the complete set of verification indicators, self-reported borrower information, loan contract information, and time fixed effects as control variables. Yet only 30% of the variation in the interest rate is explained by the model. This suggests that the interest rate does not tie closely to the observed borrower's quality, which may misinform lenders of the potential default risk and lead them to over- or underestimate the importance of some borrower characteristics.

For the three information verification indicators, except for the verification of education, none of the coefficients is significant, despite significant associations between the verification of marital status, education and employment information, and the loan default rate. Apparently, the assigned loan interest rate fails to identify the potential default risk of a loan. According to the findings in the table, the platform should lower the loan interest rate for borrowers who have verified their marital status and education, and raise the interest rates for those who have verified employment information so as to better represent the underlying default risk.

2.8 Conclusion

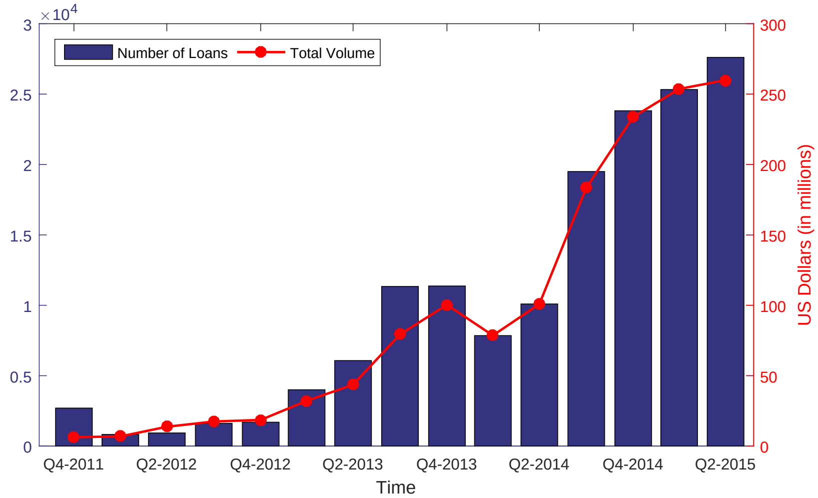
Utilizing a unique dataset from a leading Chinese P2P lending platform, this chapter studies how individual lenders make investment decisions in the online loan market. In particular, the research project attempts to investigate whether lenders make efficient funding decisions by reasonably assessing the relevance of different borrower characteristics, with an emphasis on the borrower's self-reported marital status, education, income, and working industry in addition to the corresponding information verification status. To answer these questions, I compare the impacts of each characteristic on the funding probability with the impact on the default rate.

As a unique feature of the online loan market, the dataset contains all the information of a loan and its borrower that a lender sees when making the funding decision. So, there exists no concern for omitted-variable biases in the estimation of the funding equation. However, because the loan funding process is not random, the study is challenged by sample selection issues when estimating determinants of default. To address the issue, I adopt three different methods. The first restricts attention to borrower characteristics for which there is no evidence of selection in the first stage. The second exploits unexpected variation in the number of loan posts as an instrument for loan funding in a Heckman selection model. The third follows recent work by Lee (2009) and Lechner and Melly (2010), to estimate bounds on the average treatment effect on treated.

The estimation results of the funding and default equations suggest that the P2P lenders make inefficient funding decisions. The funding probability is higher for the borrowers who have verified their employment information and self-claim to be married with high monthly income, though the corresponding default rate is not lower. On the other hand, borrowers who have verified their marital status and education level are associated with lower default risk, but they are not rewarded with better funding opportunities. Regardless of education level, verification of education does not improve funding probability but is associated with lower default risk for higher educated borrowers. For all income levels, verification of employment is extremely important for funding probability, though it is associated with higher default probability. Relative to those who self-report to be working in manufacturing, borrowers who work in finance, law, real estate, IT, and NGO are underestimated, while those who work in the construction, energy, and utility sectors are overestimated. With respect to the self-claimed education level, the lender's evaluation is reasonable. These conclusions are consistent across each of the estimation methods.

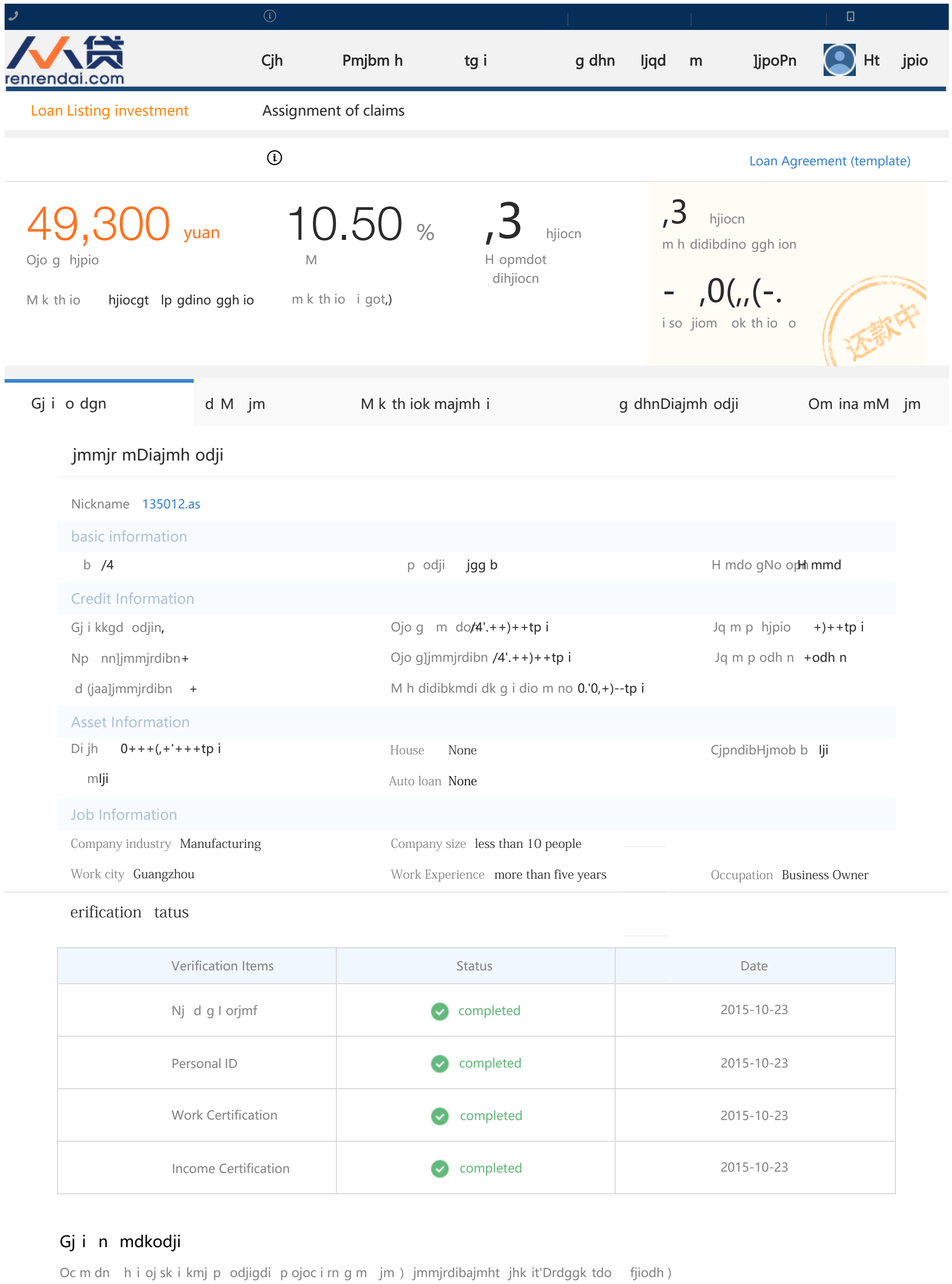
The underlying mechanism of such phenomenon is likely to be a fruitful area of research. Preliminary explanation suggests that lenders' lack of market experience may not be a primary factor. Meanwhile, the current interest rate scheme could be misleading for the lenders. The platform can potentially improve the efficiency of the market by rewriting the formula of loan interest rate to connect better with the default risks.

Figures and Tables



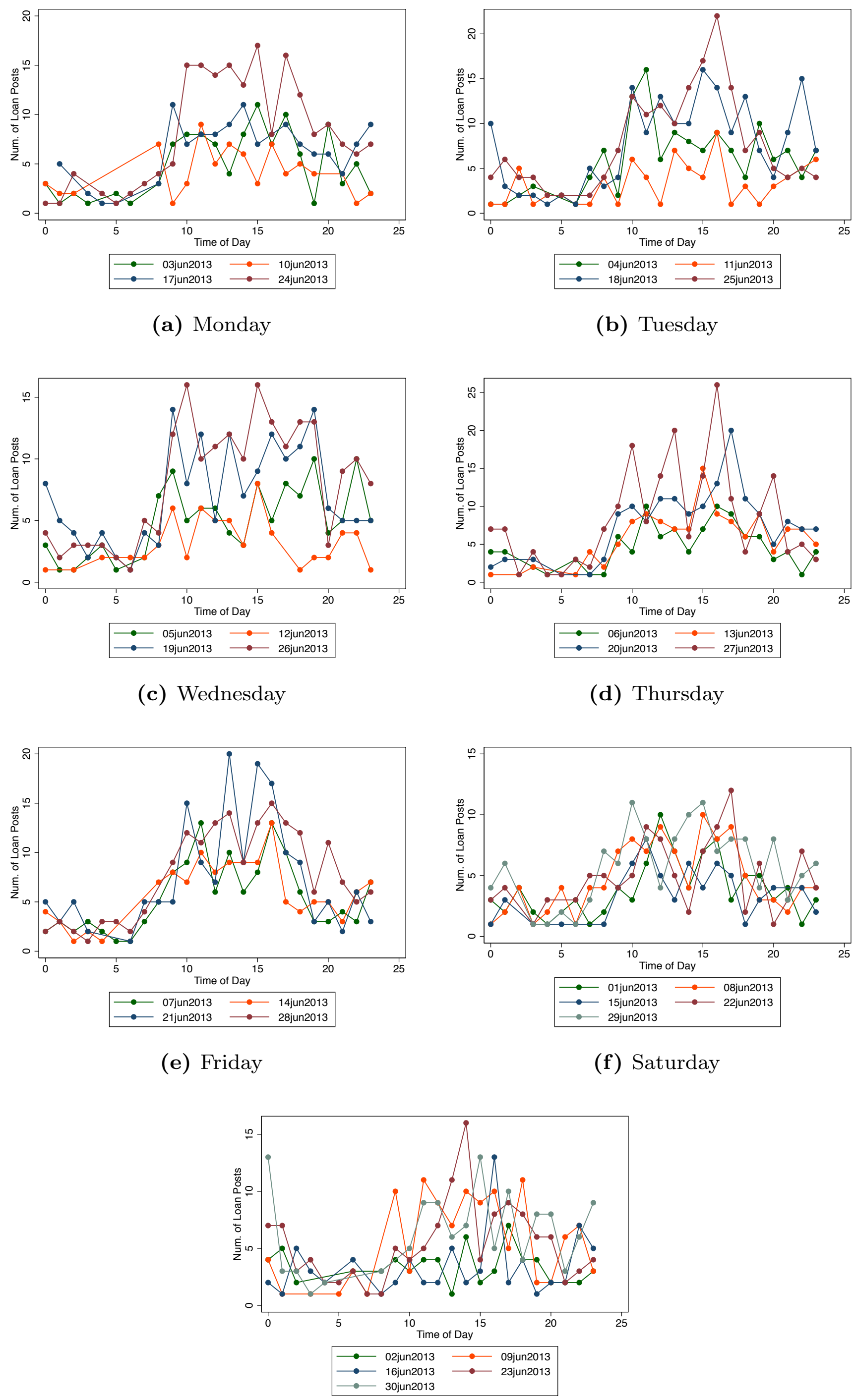
Note: Data source: quarterly and annual reports of Renrendai. All the reports are self issued by Renrendai, and available online at www.renrendai.com/about/about.action?flag=performance. The earliest report is of 2012:Q1 and the latest one is of 2015:Q2. The first observation (of Q4-2011) records the total number and volume of loans that had been made before 2012, which is reported in the quarterly report of 2012:Q1.

Figure 2.1. RRD Market Transaction



Note: The sample webpage is availale at https://www.renrendai.com/pc/loan/233732.html. And the webpage is translated into English manually.

Figure 2.2. Sample Page of Loan List on Renrendai



(a) Monday

(b) Tuesday

(c) Wednesday

(d) Thursday

(e) Friday

(f) Saturday

(g) Sunday

Figure 2.3. Number of Loans Posted in June 2013

Table 2.1. Summary Statistics

		Total	Funded	Non-Funded
Loan contract:	Interest rate	13.71%	12.68%	13.95%
		(2.76)	(1.97)	(2.86)
	Principal	$1706	$547	$1979
		(2443)	(720)	(2620)
	Term	17.88	13.84	18.84
		(10.9)	(9.6)	(11)
Verification:	VI: marital status	10.66%	16.24%	9.35%
	VI: education	10.72%	14.67%	9.79%
	VI: employment	30.09%	79.16%	18.51%
	VI: certification	2.93%	3.9%	2.71%
	VI: ID	67.40%	84.88%	63.27%
	VI: video-chat	8.5%	22.66%	5.15%
	VI: social network	5.86%	8.17%	5.31%
	VI: house ownership	13.81%	19.49%	12.47%
	VI: car ownership	10.64%	16.62%	9.23%
	VI: residence	12.61%	15.50%	11.93%
	VI: cell-phone	12.20%	15.77%	11.36%
Marital status:	Married	55.30%	61.02%	53.95%
Education:	HS and below	31.58%	24.69%	33.20%
	Associate degree	42.02%	41.38%	42.18%
	Bachelor or higher	26.40%	33.93%	24.62%
Above median income:	$800	61.83%	64.81%	61.12%
Obs.		44,765	8,547	36,218

† Interest rate is in annualized percentage rates, principal is converted to dollars using the exchange rate CNY 6.25= US$ 1, and term is in months. Standard deviations are listed in the parentheses below.

‡ Information verification indicators and self-reported information are all categorical variables.

Table 2.2. Self-claimed Information

Information	Classification
Marital status	married vs. the others*
Education	high school and below* associate degree bachelor or higher
Employment	10: agriculture, 20: construction/energy/utility, 30: manufacturing*, 40: retail/wholesale/transportation, 50: finance/law/real estate/IT, 60: education/health, 70: arts&entertainment/sports/hospitality, 80: NGO/others, 90: government. above median income* ($800)
City size	small city medium sized city large city metropolitan area
Experience	less than 1 year 1-3 years 3-5 years over 5 years
Firm size	below 10 employees 10-100 employees 100-500 employees larger than 500 employees
Property ownership	own a house vs. not own a car vs. not
Mortgage & Auto loan	with vs. without mortgage with vs. without auto loan
Borrowing purpose	personal expenditure medical expenditure wedding planning entrepreneurship/startup education/training short-term liquidity home decoration home purchase car purchase other expenditure

† The others in martial status include single, divorced and widowed. Working industries of employment are defined by the 2-digit North American Industry Classification (NAICS). City sizes are defined by the residence population according the Chinese state classification of urbanization. The population for a small city is less than 500,000; for a medium-sized city is between 500,000 to 1 million; for a large city is between 1 to 3 million; and for a metropolitan area is over 3 million.

* Indicates the baseline category for each variable studied in Section 2.4 to Section 2.6.

Table 2.3. Funding Equation: Least Squares Estimates

	(1)	(2)	(3)	(4)
VI: marital status	-0.00781	-0.00769	-0.00977	-0.00959
	(0.00833)	(0.00824)	(0.00824)	(0.00827)
VI: education	-0.00778	-0.00979	-0.00945	-0.00849
	(0.0117)	(0.0116)	(0.0114)	(0.0115)
VI: employment	0.385***	0.381***	0.379***	0.379***
	(0.0121)	(0.0118)	(0.0115)	(0.0118)
Married	0.0142***	0.0130***	0.0129***	0.0135***
	(0.00375)	(0.00352)	(0.00348)	(0.00340)
Associate degree	0.00916**	0.00880**	0.00905**	0.0103**
	(0.00394)	(0.00381)	(0.00390)	(0.00400)
Bachelor or higher	0.0279***	0.0247***	0.0254***	0.0268***
	(0.00698)	(0.00623)	(0.00622)	(0.00654)
Above median income($800)	0.0390***	0.0390***	0.0390***	0.0386***
	(0.00739)	(0.00742)	(0.00760)	(0.00764)
Agriculture	-0.00312	-0.00392	-0.00434	-0.00269
	(0.0127)	(0.0123)	(0.0119)	(0.0111)
Con, En, Ut	-0.00865	-0.00752	-0.00776	-0.00570
	(0.00636)	(0.00629)	(0.00614)	(0.00543)
Ret, whl, Tran	0.00402	0.00366	0.00308	0.00551
	(0.00493)	(0.00473)	(0.00459)	(0.00451)
Fin, law, RE, IT	-0.00317	-0.00509	-0.00458	-0.00294
	(0.00738)	(0.00726)	(0.00676)	(0.00605)
Edu, Health	-0.00293	-0.00360	-0.00379	-0.00195
	(0.0124)	(0.0121)	(0.0115)	(0.0108)
A&E, Sports, Hosp	-0.00181	-0.00112	-0.00379	-0.00102
	(0.00601)	(0.00591)	(0.00617)	(0.00597)
NGO, others	-0.00340	-0.00459	-0.00508	-0.00282
	(0.00625)	(0.00644)	(0.00632)	(0.00585)
Gov	0.0168	0.0164	0.0155	0.0181
	(0.0165)	(0.0163)	(0.0155)	(0.0156)
Loan interest rate	No	Continuous	Categorical	Interacted w/ Mon. FE
Controls	✓	✓	✓	✓
R^2	0.359	0.366	0.373	0.383
Obs.	44,765	44,765	44,765	44,765

† Control variables include all the other information verification indicators and self-claimed borrower information. The baseline of education is high school and below and that of working industry of employment is manufacturing.
‡ Loan interest rate is the annualized percentage rate of loan interest rate which is excluded in column (1) and included as a continuous variable in column (2) and as a discrete variable in column (3). Column (4) includes the interaction term between loan interest rate and monthly fixed effects.
§ The standard errors are clustered by the borrower's quality.
* Indicates statistically significant at the 0.1 level.
** Indicates statistically significant at the 0.5 level.
*** Indicates statistically significant at the 0.01 level.

Table 2.4. Default Equation: Least Squares Estimates

	(1)	(2)	(3)	(4)
VI: marital status	-0.0383***	-0.0382***	-0.0404***	-0.0385***
	(0.0129)	(0.0130)	(0.0131)	(0.0131)
VI: education	-0.0572***	-0.0580***	-0.0577***	-0.0626***
	(0.0139)	(0.0135)	(0.0133)	(0.0126)
VI: employment	0.135***	0.133***	0.134***	0.135***
	(0.0425)	(0.0431)	(0.0432)	(0.0436)
Married	-0.00889	-0.00883	-0.00843	-0.00737
	(0.00824)	(0.00846)	(0.00813)	(0.00787)
Associate degree	-0.0264**	-0.0258**	-0.0265**	-0.0257**
	(0.0131)	(0.0125)	(0.0122)	(0.0127)
Bachelor or higher	-0.108***	-0.104***	-0.105***	-0.102***
	(0.0223)	(0.0215)	(0.0214)	(0.0220)
Above median income($800)	0.0267**	0.0273**	0.0272**	0.0267**
	(0.0123)	(0.0117)	(0.0119)	(0.0118)
Agriculture	0.0583*	0.0553	0.0544	0.0670*
	(0.0350)	(0.0358)	(0.0354)	(0.0348)
Con, En, Ut	0.0373*	0.0366*	0.0348*	0.0327*
	(0.0196)	(0.0188)	(0.0189)	(0.0169)
Ret, whl, Tran	-0.00332	-0.00263	-0.00376	-0.00414
	(0.0105)	(0.0102)	(0.0103)	(0.0102)
Fin, law, RE, IT	-0.0617***	-0.0600***	-0.0604***	-0.0587***
	(0.0152)	(0.0154)	(0.0150)	(0.0141)
Edu, Health	-0.0183	-0.0168	-0.0176	-0.0190
	(0.0121)	(0.0122)	(0.0121)	(0.0119)
A&E, Sports, Hosp	-0.00457	-0.00605	-0.00906	-0.0110
	(0.0151)	(0.0153)	(0.0146)	(0.0134)
NGO, others	-0.0338***	-0.0313***	-0.0335***	-0.0348***
	(0.0118)	(0.0115)	(0.0118)	(0.0114)
Gov	-0.0114	-0.0114	-0.0137	-0.0147
	(0.0126)	(0.0121)	(0.0122)	(0.0122)
Loan interest rate	No	Continuous	Categorical	Interacted w/ Mon. FE
Controls	✓	✓	✓	✓
R^2	0.125	0.131	0.135	0.158
Obs.	8,547	8,547	8,547	8,547

† Control variables include all the other information verification indicators and self-claimed borrower information. The baseline of education is high school and below and that of working industry of employment is manufacturing.

‡ Loan interest rate is the annualized percentage rate of loan interest rate which is excluded in column (1) and included as a continuous variable in column (2) and as a discrete variable in column (3). Column (4) includes the interaction term between loan interest rate and monthly fixed effects.

§ The standard errors are clustered by the borrower's quality.

* Indicates statistically significant at the 0.1 level.

** Indicates statistically significant at the 0.5 level.

*** Indicates statistically significant at the 0.01 level.

Table 2.5. Inference on the Coefficient Estimates Comparison

		Default Eq.		
		$\tilde{\beta}_\kappa > 0$	$\tilde{\beta}_\kappa < 0$	$\tilde{\beta}_\kappa = 0$
	$\alpha_\kappa = 0$	Overestimated	Underestimated	Reasonable
Funding Eq.	$\alpha_\kappa > 0$	Overestimated	Reasonable	Overestimated
	$\alpha_\kappa < 0$	Reasonable	Underestimated	Underestimated

† α_κ can be considered as either $\alpha_{1\kappa}$ or $\alpha_{2\kappa}^j$ and $\tilde{\beta}_\kappa$ as either $\tilde{\beta}_{1\kappa}$ or $\tilde{\beta}_{2\kappa}^j$ for all $j \in J(\kappa)$ and κ in Equation (2.1) and (2.2).

‡ The overestimation (or underestimation) is defined as the lender's subjective expectation predicts a default risk lower (or higher) than that is predicted by the true model.

Table 2.6. Effect of Number of Loan Posts on Funding Rate

	(1)	(2)	(3)
Number of loan posts	-0.00192***	-0.00200***	-0.00198***
	(0.000339)	(0.000273)	(0.000273)
Time fixed effects:			
Month	Yes	Yes	Yes
Day in a week	Yes	Yes	Yes
Hour in a day	Yes	Yes	Yes
Loan interest rate	Continuous	Continuous	Continuous
Controls	No	Yes	w/ interaction
R^2	0.017	0.368	0.369
Obs.	44,765	44,765	44,765

† Number of loan posts is the total number of loans posted within the same hour. Control variables include all the information verification indicators and self-claimed borrower information in column (2) and the interaction terms between the information verification indicators and self-claimed borrower information in column (3).

§ The standard errors are clustered by the borrower's quality.

* Indicates statistically significant at the 0.1 level.

** Indicates statistically significant at the 0.5 level.

*** Indicates statistically significant at the 0.01 level.

Table 2.7. Relationship between Number of Loan Posts and Borrower Characteristics

	(1)	(2)
VI: marital Status	0.144	-0.343
	(0.0996)	(0.605)
VI: education	0.0246	-1.481
	(0.0930)	(2.334)
VI: employment	-0.0420	-0.113
	(0.195)	(0.220)
Married	0.0158	0.507
	(0.0608)	(0.607)
Associate degree	-0.0863	1.383
	(0.0649)	(2.337)
Bachelor or higher	-0.108	1.438
	(0.0798)	(2.336)
Above median income($800)	0.0194	0.0877
	(0.0623)	(0.109)
Time fixed effects:		
Month	Yes	Yes
Day in a week	Yes	Yes
Hour in a day	Yes	Yes
Loan interest rate	Continuous	Continuous
Controls	Yes	w/ interaction
R^2	0.683	0.822
Obs.	44,765	44,765

† Control variables include all the information verification indicators and self-claimed borrower information in column (1) and the interaction terms between the information verification indicators and self-claimed borrower information in column (2).
§ The standard errors are clustered by the borrower's quality.
* Indicates statistically significant at the 0.1 level.
** Indicates statistically significant at the 0.5 level.
*** Indicates statistically significant at the 0.01 level.

Table 2.8. Funding and Default Equations: Instrument Variable Estimates

	Funding Eq.	Default Eq.
VI: marital status	-0.00769	-0.0360***
	(0.00824)	(0.0134)
VI: education	-0.00979	-0.0571***
	(0.0116)	(0.0142)
VI: employment	0.381***	0.102**
	(0.0118)	(0.0423)
Married	0.0130***	-0.00927
	(0.00352)	(0.00946)
Associate degree	0.00880**	-0.0275**
	(0.00381)	(0.0123)
Bachelor or higher	0.0247***	-0.105***
	(0.00623)	(0.0199)
Above median income($800)	0.0390***	0.0220
	(0.00742)	(0.0176)
Agriculture	-0.00392	0.0573
	(0.0123)	(0.0357)
Con, En, Ut	-0.00752	0.0358**
	(0.00629)	(0.0172)
Ret, whl, Tran	0.00366	-0.00323
	(0.00473)	(0.0103)
Fin, law, RE, IT	-0.00509	-0.0600***
	(0.00726)	(0.0164)
Edu, Health	-0.00360	-0.0181
	(0.0121)	(0.0122)
A&E, Sports, Hosp	-0.00112	-0.00846
	(0.00591)	(0.0154)
NGO, others	-0.00459	-0.0322***
	(0.00644)	(0.0125)
Gov	0.0164	-0.0146
	(0.0163)	(0.0120)
Number of loan posts	-	-0.0106***
	-	(0.00180)
Loan interest rate	Continuous	Continuous
Controls	Yes	Yes
Time fixed effects:		
Month	Yes	Yes
Day in a week	No	Yes
Hour in a day	No	Yes
Obs.	44,765	44,765

† The coefficients of funding equation are estimated using ordinary least squares. The coefficients of default equation are estiamted by Heckman selecion model using maximum likelihood method. The number of loan posts is included as the exogenous variation in the sample selection. Control variables include all the other information verification indicators and self-claimed borrower information. The standard errors are clustered by the borrower's quality.

* Indicates statistically significant at the 0.1 level.

** Indicates statistically significant at the 0.5 level.

*** Indicates statistically significant at the 0.01 level.

Table 2.9. Funding and Default Equations: with Interaction Terms

	Funding Eq.	Default Eq.
VI: marital status	0.0523	-0.0589
	(0.0338)	(0.0529)
Married	0.0132***	-0.0102
	(0.00342)	(0.00957)
VI: marital status × Married	-0.0610*	0.0261
	(0.0341)	(0.0544)
VI: education	-0.0210	0.0999***
	(0.0220)	(0.0332)
Associate degree	0.00494	-0.0158*
	(0.00363)	(0.00923)
Bachelor or higher	0.0221***	-0.0925***
	(0.00459)	(0.0157)
VI: education × Associate degree	0.00357	-0.146***
	(0.0208)	(0.0380)
VI: education × Bachelor or higher	0.0219	-0.166***
	(0.0206)	(0.0315)
VI: employment	0.408***	0.0901**
	(0.00565)	(0.0370)
Above median income($800)	0.0517***	0.0198
	(0.00398)	(0.0162)
VI: employment × Above median income($800)	-0.0429***	-0.00210
	(0.00676)	(0.0154)
Loan interest rate	Yes	Yes
Controls	Yes	Yes
Time fixed effects:		
Month	Yes	Yes
Day in a week	No	Yes
Hour in a day	No	Yes
Obs.	44,765	44,765

† Funding equation is estiamted using least squares. Default equation is estiamted by Heckman selecion model using maximum likelihood method. The number of loan posts is included as the exogenous variation in the sample selection. Control variables include all the other information verification indicators and self-claimed borrower information. The standard errors are clustered by the borrower's quality.

‡ Standard errors are in the paranthese and P-values are in the bracket.

* Indicates statistically significant at the 0.1 level.

** Indicates statistically significant at the 0.5 level.

*** Indicates statistically significant at the 0.01 level.

Table 2.10. Inference on the Bounded ATET

		Default Eq.		
		$\tilde{\beta}_{\kappa}^{LB} > 0$	$\tilde{\beta}_{\kappa}^{UB} < 0$	$0 \in \left[\tilde{\beta}_{\kappa}^{LB}, \tilde{\beta}_{\kappa}^{UB}\right]$
Funding Eq.	$\alpha_{\kappa} > 0$	Overestimated	Reasonable	Inconclusive
	$\alpha_{\kappa} < 0$	Reasonable	Underestimated	Inconclusive

† α_{κ} can be considered as either $\alpha_{1\kappa}$ or $\alpha_{2\kappa}^{j}$ and $\tilde{\beta}_{\kappa}$ as either $\tilde{\beta}_{1\kappa}$ or $\tilde{\beta}_{2\kappa}^{j}$ for all $j \in J(\kappa)$ and κ in Equation (2.1) and (2.2). The superscript of LB and UB denote the lower and upper bounds estimated following Lee (2009).

‡ The overestimation (or underestimation) is defined as the lender's subjective expectation predicts a default risk lower (or higher) than that is predicted by the true model.

Table 2.11. Bounds on the Average Treatment Effects on the Treated

	Funding Eq.	Default Eq.		
		Upper Bd.	Lower Bd.	95% CI
VI: employment	0.381*** (0.0118)	-0.017 (0.0446)	-0.187 (0.0402)	[-0.253, 0.057]
Married	0.0130*** (0.00352)	-0.031 (0.013)	-0.082 (0.024)	[-0.08, -0.009]
Associate degree	0.00880** (0.00381)	-0.054 (0.018)	-0.130 (0.028)	[-0.185, -0.019]
Bachelor or higher	0.0247*** (0.00623)	-0.188 (0.017)	-0.317 (0.020)	[-0.360, -0.160]
Above median income($800)	0.0390*** (0.00742)	0.026 (0.014)	0.026 (0.023)	[-0.019, 0.054]

† The boundaries are estimated separately for each variable in the table. The number of observations used in the estimation is 699 for the verification of employment, 6341 for the self-reported marital status, 10313 for the self-reported associate degree, 12886 for the self-reported bachelor degree or higher, and 13030 for the self-reported income level. The standard errors and the confidence intervals are calculated following Imbens and Manski (2004) and Lee (2009). The details can be found in Appendix 2.B.

‡ The least squares estimation of the funding equation in column (2) of Table 2.3 are included for inference.

Table 2.12. Bounds on the Average Treatment Effects: Using the Entire Sample

	Effect 95% Confidence Interval	
	Lower boundary	Upper boundary
VI: employment	-0.1646	0.8326
Married	-0.2106	0.0557
Associate degree	0.0199	0.0982
Bachelor or higher	-0.2905	-0.0220
Above median income($800)	-0.1397	0.0408
Funded Obs.	8,547	
Total Obs.	44,765	

† The boundaries are estimated separately for each variable in the table using the entire sample of observations. The confidence intervals are calculated following Imbens and Manski (2004) and Lee (2009).

Table 2.13. Effects of Information Verification on Funding, Default and Interest Rate

	Funding Eq.			Default Eq.	Loan Interest Rate
	Full sample	1^{st} half	2^{nd} half	Full sample	
VI: marital status	-0.00769	-0.0147	-0.00131	-0.0360***	7.64e-05
	(0.00824)	(0.0124)	(0.00960)	(0.0134)	(0.000613)
VI: education	-0.00979	0.0136	-0.0138	-0.0571***	-0.00138**
	(0.0116)	(0.0126)	(0.0117)	(0.0142)	(0.000576)
VI: employment	0.381***	0.127***	0.411***	0.102***	-0.000693
	(0.0118)	(0.0196)	(0.0118)	(0.0423)	(0.00127)
Loan interest rate	Continuous	Continuous	Continuous	Continuous	-
Controls	Yes	Yes	Yes	Yes	Yes
R^2	0.366	0.454	0.412	-	0.308
Obs.	44,765	9,992	34,773	44,765	44,765

† 1^{st} half spans from June 2013 to March 2014, and 2^{nd} half from April to December 2014. The coefficients of funding equations are estimated using ordinary least squaes. The coefficients of default equation in full sample are estimated by Heckman selecton model using maximum likelihood method as in section 2.5. Control variables include the information verification indicators and self-claimed borrower information. The standard errors are clustered by the borrower's quality.

‡ The column of loan interest rate reports the results of the least squares estimation of Equation (2.7) with the same explantory variables as in the Heckman selecion model in Section 2.5. The standard errors are clustered by the borrower's quality.

* Indicates statistically significant at the 0.1 level.

** Indicates statistically significant at the 0.5 level.

*** Indicates statistically significant at the 0.01 level.

References

Agrawal, Ajay, Christian Catalini, and Avi Goldfarb. 2015. “Crowdfunding: Geography, Social Networks, and the Timing of Investment Decisions.” *Journal of Economics and Management Strategy*, 24(2): 253–274.

Antweiler, Werner, and Murray Z. Frank. 2004. “Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards.” *The Journal of Finance*, LIX(3): 1259–1294.

Barasinska, Nataliya, and Dorothea Schafer. 2014. “Is Crowdfunding Different? Evidence on the Relation Between Gender and Funding Success from a German Peer-to-Peer Lending Platform.” *German Economic Review*, 15(4): 436–452.

Burtch, Gordon, Anindya Ghose, and Sunil Wattal. 2014. “Cultural Differences and Geography as Determinants of Online Prosocial Lending.” *MIS Quarterly*, 38(3): 773–794.

Chen, Baizhu, and Wei Xu. 2015. “Credit Insurance in the Peer-to-Peer Lending Market: Evidence from China.” *Working Paper.*

Chen, Dongyu, and Zhangxi Lin. 2014. “Rational or Irrational Herding in Online Microloan Markets: Evidence from China.” *SSRN Working Paper, 2425047.*

Dorfleitner, Gregor, Christopher Priberny, Stephanie Schuster, Johannes Stoiber, Martina Weber, Ivan de Castro, and Julia Kammler. 2016. “Description-Text Related Soft Information in Peer-to-Peer Lending - Evidence from Two Leading European Platforms.” *Journal of Banking and Finance*, 64: 169–187.

Duarte, Jefferson, Stephan Siegel, and Lance Young. 2012. “Trust and Credit: The Role of Appearance in Peer-to-Peer Lending.” *Review of Financial Studies*, 25: 2455–2484.

Engelberg, Joseph. 2008. “Costly Information Processing: Evidence from Earnings Announcements.” *Working paper, Northwestern University.*

Freedman, Seth, and Ginger Zhe Jin. 2008. “Do Social Networks Solve Information Problems for Peer-to-Peer Lending? Evidence from Prosper.com.” *SSRN Working Paper, 1936057.*

Herzenstein, Michal, Scott Sonenshein, and Utpal M. Dholakia. 2011. "Tell Me a Good Story and I May Lend You My Money: The Role of Narratives in Peer-to-Peer Lending Decisions." *Journal of Marketing Research*, XLVIII: S138–S149.

Hollingshead, August B. 1975. *Four Factor Index of Social Status.* New Haven, Conn. : Yale University, Dept. of Sociology.

Imbens, Guido W., and Charles F. Manski. 2004. "Confidence Intervals for Partially Identified Parameters." *Econometrica*, 72: 1845–1857.

Iyer, Rajkamal, Asim Ijaz Khwaja, Erzo F. P. Luttmer, and Kelly Shue. 2015. "Screening Peers Softly: Inferring the Quality of Small Borrowers." *Management Science*, 62(6): 1554–1557.

Kim, Moshe, Jordi Surroca, and Josep A. Tribo. 2014. "Impact of Ethical Behavior on Syndicated Loan Rates." *Journal of Banking and Finance*, 38: 122–144.

Lechner, Michael, and Blaise Melly. 2010. "Partial Identification of Wage Effects of Training Programs." *SSRN Working Paper, 1596715.*

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*, 76: 1071–1102.

Li, Feng. 2008. "Annual Report Readability, Current Earnings, and Earnings Persistence." *Journal of Accounting and Economics*, 45: 221–247.

Lin, Mingfeng, and Siva Viswanathan. 2016. "Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market." *Management Science*, 62(5): 1393–1414.

Lin, Mingfeng, Nagpurnanand R. Prabhala, and Siva Viswanathan. 2013. "Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending." *Management Science*, 59(1): 17–35.

Liu, De, Daniel J. Brass, Yong Lu, and Dongyu Chen. 2015. "Friendships in Online Peer-to-Peer Lending: Pipes, Prisms, and Relational Herding." *MIS Quarterly*, 39(3): 729–742.

Marom, Dan, Alicia Robb, and Orly Sade. 2016. "Gender Dynamics in Crowdfunding (Kickstarter): Evidence on Entrepreneurs, Investors, Deals, and Taste-Based Discrimination." *SSRN Working Paper.*

Michels, Jeremy. 2012. "Do Unverifiable Disclosures Matter? Evidence from Peer-to-Peer Lending." *The Accounting Review*, 87(4): 1385–1413.

Pope, Devin G., and Justin R. Sydnor. 2011. "What's in a Picture? Evidence of Discrimination from Prosper.com." *The Journal of Human Resources*, 46(1).

Ravina, Enrichetta. 2012. "Love and Loans: The Effect of Beauty and Personal Characteristics in Credit Markets." *SSRN Working Paper, 1101647.*

Sonenshein, Scott, Michal Herzenstein, and Utpal M. Dholakia. 2011. "How Accounts Shape Lending Decisions Through Fostering Perceived Trustworthiness." *Organizational Behavior and Human Decision Processes*, 115: 69–84.

Tetlock, Paul C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *The Journal of Finance*, LXII(3): 1139–1168.

Tetlock, Paul C., Maytal Saar-Tsecuansky, and Sofus Macskassy. 2008. "More Than Words: Quantifying Language to Measure Firms' Fundmentals." *The Journal of FInance*, LXIII(3): 1437–1467.

Zhang, Juanjuan, and Peng Liu. 2012. "Rational Herding in Microloan Markets." *Management Science*, 58(5): 892–912.

Appendices

2.A Soft Information From Loan Descriptions

2.A.1 Top 100 Most Common Words and Phrases

When applying for P2P loans, the borrowers are required to describe the purpose and the use of their loan requests. The top 100 words and phrases are listed as follows ordered by their frequencies in loan description: money, transition, myself, borrow, repay, need, hope, firm, thanks, decoration, everyone, credit, work, loan, support, gain, operation, used for, stable, current, short-term, me, on-time, therefore, urgent, can, monthly, friend, business, personal, income, good, house, because, expend, help, apply, capability, invest, creditable, improvement, momentarily, startup, consumption, guarantee, because, purchase, one, no, platform, majorly, one, thousand yuan, marital status, preparation, more, job, for sure, first-time, through, family, sales, time, approximately, company, project, already, bank, corporate, end of year, problem, recently, production, merchandise, shop, immediately, this, business, doing, end of year, problem, recently, production, honest, merchandise, shop, immediately, this, business, doing, in stock, property, record, increase, our, equipment, pressure, thanks, acquire, rise, safe your heart, comparison, some, life, development, successes.

2.A.2 Positive Words Among Top 100 Words and Phrases

Positive words defined according to the explanation in Xin Hua Dictionary and Xin Hua 08 Chinese-English Financial Dictionary. The positive words from the top 100 words and phrases in the P2P loan description are stable, good, expand, capability, creditable, guarantee, development, honest, increase, improvement, safe your heart, successes.

2.B Calculation of the Standard Errors and Confidence Intervals for the Bounded Average Treatment Effects

Let's denote N_C =number of control observations, N_{FC} = number of funded control observations, N_T = number of treatment observations, and N_{FT} =number of funded treatment observations. And let $\pi_T = N_{FT}/N_T$ as the proportion of funded loans in treatment and $\pi_C = N_{FC}/N_C$ the proportion of funded loans in control. $\bar{D}_{FC} \equiv E(D_{FC})$ is the average default rate for the funded loans in control while $\bar{D}_{FT}(q) = E[D_{FT}|D_{FT} > D_{FT}(q)]$ and $\tilde{D}_{FT}(q) = E[D_{FT}|D_{FT} < D_{FT}(q)]$ are the trimmed default rate averages for the funded loans in treatment, where $D_{FT}(q)$ is the qth quantile of the default rates for the funded loans in treatment.

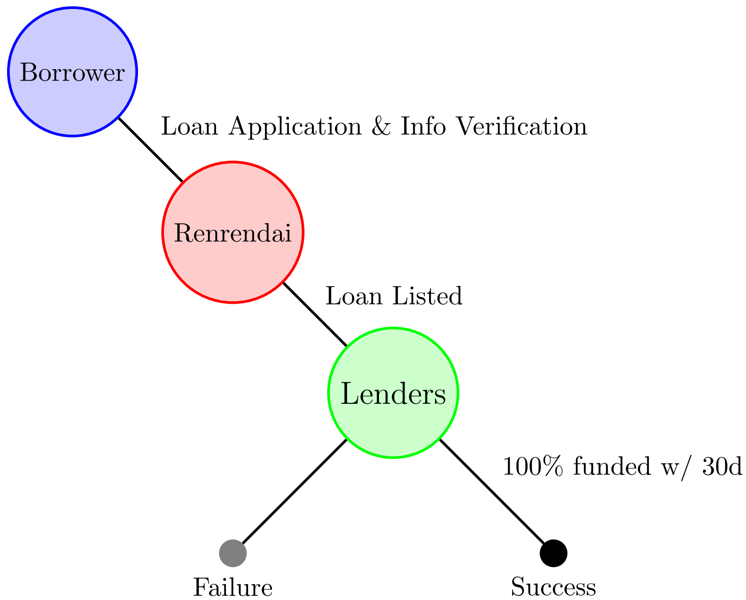
Following Lee (2009), the standard error of the upper bound, SE^{UB}, includes three components. Component 1 is the usual standard error of the mean, using the trimmed sample. Component 2 is the square root of $(1/N_{FC}) \times (q/(1-q)) \times \left[\bar{D}_{FT}(q) - D_{FT}(q)\right]^2$. Component 3 is the square root of $\left[\left(\bar{D}_{FT}(q) - D_{FT}(q)\right)/(1-q)\right]^2 \times Var(q)$, where $Var(q) = (1-q)^2 \times \{(1/N_T) \times [(1-\pi_T)/\pi_T] + (1/N_C) \times [(1-\pi_C)/\pi_C]\}$. The "total" standard error is the square root of the sum the squared components. The standard error of the lower bound, SE^{LB}, can be calculated analogously where $D_{FT}(q)$ and $\bar{D}_{FT}(q)$ are replaced by $D_{FT}(1-q)$ and $\tilde{D}_{FT}(1-q)$ respectively.

The confidence intervals are computed using the method proposed in Imbens and Manski (2004). The 95% confidence interval is given by

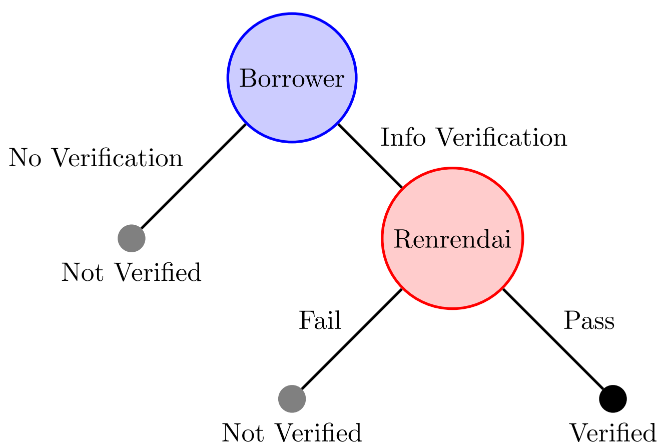
$$\left[\left(\tilde{D}_{FT}(1-q) - \bar{D}_{FC}\right) - f_{95} \times SE^{LB}, \left(\bar{D}_{FT}(q) - \bar{D}_{FC}\right) + f_{95} \times SE^{UB}\right],$$

where f_{95} satisfies $\Phi\left(f_{95} + \left(\bar{D}_{FT}(q) - \tilde{D}_{FT}(1-q)\right)/\max\left(SE^{UB}, SE^{LB}\right)\right) - \Phi(-f_{95}) = 0.95$ with Φ standing for the c.d.f. of standard normal distribution.

2.C Complementary Figures and Tables



(a) Borrowing Procedure



(b) Procedure of Borrower Information Verification

Figure 2.C.1. Borrowing at Renrendai

Table 2.C.1. Average Interest Rate and Number of Information Verification

Num. of VI	Avg. APR	(Std.)	Obs.
1	13.71%	(2.85)	8487
2	13.77%	(2.75)	13958
3	14.07%	(3.03)	5116
4	13.50%	(2.49)	7618
5	13.62%	(2.66)	4058
6	13.65%	(2.70)	2323
7	13.72%	(2.71)	1555
8	13.47%	(2.83)	765
9	13.52%	(2.82)	495
10	13.06%	(2.83)	315
11	13.36%	(2.19)	75
Overall	13.71%	(2.76)	44765

† Interest rate is annualized percentage rate.

Table 2.C.2. Average Interest Rate and Verified Information

Verified Information	Avg. APR	(Stdv.)	Obs.
Marital status	13.76%	(2.94)	4774
Education	13.56%	(2.66)	4799
Employment	13.39%	(2.44)	13471
Certification	13.70%	(2.83)	1313
ID	13.70%	(2.75)	44349
Video-chat	13.78%	(2.82)	3803
Social network	13.57%	(2.89)	2622
House ownership	13.56%	(2.68)	6184
Car ownership	13.50%	(2.68)	4765
Residence	14.35%	(3.16)	5644
Cell-phone	14.09%	(3.11)	5463
ID & Employment	13.40%	(2.43)	12918

† Interest rate is annualized percentage rate.

Table 2.C.3. Funding Equation with Control Interaction Terms: Least Squares Estimates

	(1)	(2)	(3)	(4)
VI: marital status	0.0584	0.0534	0.0541	0.0595
	(0.0492)	(0.0487)	(0.0497)	(0.0492)
VI: education	-0.00560	-0.000220	0.00119	0.00396
	(0.0277)	(0.0274)	(0.0272)	(0.0273)
VI: employment	0.411***	0.408***	0.406***	0.407***
	(0.00983)	(0.00894)	(0.00869)	(0.00876)
Married	-0.0524	-0.0484	-0.0513	-0.0560
	(0.0494)	(0.0488)	(0.0498)	(0.0493)
Associate degree	0.00797**	0.00752**	0.00772**	0.00893**
	(0.00386)	(0.00373)	(0.00384)	(0.00393)
Bachelor or higher	0.0621***	0.0545***	0.0554***	0.0554***
	(0.0198)	(0.0199)	(0.0203)	(0.0204)
Above median income($800)	0.00797	0.00786	0.00765	0.00641
	(0.0105)	(0.0108)	(0.0107)	(0.0112)
Agri	-0.00387	-0.00460	-0.00507	-0.00333
	(0.0126)	(0.0122)	(0.0118)	(0.0110)
Con, En, Ut	-0.00970	-0.00858	-0.00882	-0.00673
	(0.00648)	(0.00644)	(0.00627)	(0.00551)
Ret, whl, Tran	0.00381	0.00343	0.00284	0.00529
	(0.00483)	(0.00465)	(0.00452)	(0.00443)
Fin, law, RE, IT	-0.00362	-0.00544	-0.00494	-0.00325
	(0.00745)	(0.00731)	(0.00681)	(0.00608)
Edu, Health	-0.00346	-0.00403	-0.00424	-0.00234
	(0.0124)	(0.0121)	(0.0115)	(0.0108)
A&E, Sports, Hosp	-0.00169	-0.00100	-0.00364	-0.000789
	(0.00591)	(0.00581)	(0.00607)	(0.00586)
NGO, others	-0.00388	-0.00504	-0.00552	-0.00321
	(0.00622)	(0.00641)	(0.00629)	(0.00581)
Gov	0.0141	0.0138	0.0129	0.0155
	(0.0171)	(0.0169)	(0.0162)	(0.0163)
Loan interest rate	No	Continuous	Categorical	Interacted w/ Mon. FE
Controls (w/ interaction)	Yes	Yes	Yes	Yes
R^2	0.360	0.367	0.373	0.383
Obs.	44,765	44,765	44,765	44,765

† Control variables include information verification indicators, self-claimed borrower information and the interaction terms between these two.
‡ Loan interest rate is the annualized percentage rate of loan interest rate which is excluded in column (1) and included as a continuous variable in column (2) and as a discrete variable in column (3). Column (4) includes the interaction term between loan interest rate and monthly fixed effects.
§ The standard errors are clustered by the borrower's quality.
* Indicates statistically significant at the 0.1 level.
** Indicates statistically significant at the 0.5 level.
*** Indicates statistically significant at the 0.01 level.

Table 2.C.4. Default Equation with Control Interaction Terms: Least Squares Estimates

	(1)	(2)	(3)	(4)
VI: marital status	-0.0640	-0.0609	-0.0609	-0.0667
	(0.0532)	(0.0535)	(0.0536)	(0.0600)
VI: education	0.0164	0.0115	0.00815	0.00287
	(0.0326)	(0.0328)	(0.0319)	(0.0325)
VI: employment	0.142***	0.141***	0.143***	0.145***
	(0.0449)	(0.0461)	(0.0458)	(0.0460)
Married	0.0182	0.0152	0.0133	0.0223
	(0.0538)	(0.0541)	(0.0542)	(0.0601)
Associate degree	-0.0260**	-0.0252**	-0.0256**	-0.0244*
	(0.0127)	(0.0120)	(0.0118)	(0.0124)
Bachelor or higher	-0.142***	-0.134***	-0.135***	-0.138***
	(0.0450)	(0.0448)	(0.0446)	(0.0427)
Above median income($800)	0.0257*	0.0259*	0.0256*	0.0247*
	(0.0146)	(0.0140)	(0.0142)	(0.0144)
Agri	0.0610*	0.0580	0.0571	0.0702**
	(0.0349)	(0.0358)	(0.0354)	(0.0346)
Con, En, Ut	0.0377*	0.0369*	0.0351*	0.0331*
	(0.0201)	(0.0193)	(0.0194)	(0.0173)
Ret, whl, Tran	-0.00332	-0.00263	-0.00373	-0.00408
	(0.0106)	(0.0103)	(0.0104)	(0.0102)
Fin, law, RE, IT	-0.0604***	-0.0588***	-0.0593***	-0.0574***
	(0.0150)	(0.0152)	(0.0149)	(0.0140)
Edu, Health	-0.0167	-0.0153	-0.0160	-0.0172
	(0.0123)	(0.0123)	(0.0123)	(0.0120)
A&E, Sports, Hosp	-0.00497	-0.00647	-0.00953	-0.0117
	(0.0151)	(0.0154)	(0.0146)	(0.0135)
NGO, others	-0.0331***	-0.0306***	-0.0328***	-0.0341***
	(0.0119)	(0.0116)	(0.0118)	(0.0114)
Gov	-0.00974	-0.00998	-0.0123	-0.0132
	(0.0126)	(0.0121)	(0.0123)	(0.0123)
Loan interest rate	No	Continuous	Categorical	Interacted w/ Mon. FE
Controls (w/ interaction)	✓	✓	✓	✓
R^2	0.125	0.132	0.135	0.158
Obs.	8,547	8,547	8,547	8,547

† Control variables include information verification indicators, self-claimed borrower information and the interaction terms between these two.

‡ Loan interest rate is the annualized percentage rate of loan interest rate which is excluded in column (1) and included as a continuous variable in column (2) and as a discrete variable in column (3). Column (4) includes the interaction term between loan interest rate and monthly fixed effects.

§ The standard errors are clustered by the borrower's quality.

* Indicates statistically significant at the 0.1 level.

** Indicates statistically significant at the 0.5 level.

*** Indicates statistically significant at the 0.01 level.

Table 2.C.5. Funding Equation without Control Variables: Least Squares Estimates

	(1)	(2)	(3)	(4)
VI: marital status	0.00795	0.00795	0.00399	-0.000392
	(0.0134)	(0.0134)	(0.0131)	(0.0125)
VI: education	-0.00355	-0.00355	-0.00602	-0.00314
	(0.0129)	(0.0129)	(0.0125)	(0.0123)
VI: employment	0.442***	0.442***	0.431***	0.425***
	(0.0117)	(0.0117)	(0.0128)	(0.0137)
Married	0.0180***	0.0180***	0.0148***	0.0135**
	(0.00506)	(0.00506)	(0.00549)	(0.00569)
Associate degree	0.0179***	0.0179***	0.0148***	0.0160***
	(0.00492)	(0.00492)	(0.00503)	(0.00442)
Bachelor or higher	0.0335***	0.0335***	0.0248***	0.0262***
	(0.0102)	(0.0102)	(0.00894)	(0.00830)
Above median income($800)	0.0104**	0.0104**	0.0125***	0.00885**
	(0.00488)	(0.00488)	(0.00450)	(0.00448)
Agri	-0.0378**	-0.0378**	-0.0347**	-0.0314**
	(0.0187)	(0.0187)	(0.0172)	(0.0146)
Con, En, Ut	-0.00792	-0.00792	-0.00445	-0.00169
	(0.00733)	(0.00733)	(0.00744)	(0.00581)
Ret, whl, Tran	-0.0134***	-0.0134***	-0.0135***	-0.00997**
	(0.00498)	(0.00498)	(0.00507)	(0.00441)
Fin, law, RE, IT	-0.0102	-0.0102	-0.0127	-0.00760
	(0.00863)	(0.00863)	(0.00816)	(0.00602)
Edu, Health	-0.00829	-0.00829	-0.00406	0.000552
	(0.0138)	(0.0138)	(0.0124)	(0.0108)
A&E, Sports, Hosp	-0.0102	-0.0102	-0.0126*	-0.00808
	(0.00680)	(0.00680)	(0.00746)	(0.00723)
NGO, others	-0.0249***	-0.0249***	-0.0235***	-0.0229***
	(0.00688)	(0.00688)	(0.00738)	(0.00655)
Gov	0.0230*	0.0230*	0.0214	0.0232
	(0.0139)	(0.0139)	(0.0135)	(0.0143)
Loan interest rate	No	Continuous	Categorical	Interacted w/ Mon. FE
No control	✓	✓	✓	✓
R^2	0.275	0.275	0.312	0.329
Obs.	44,765	44,765	44,765	44,765

† Loan interest rate is the annualized percentage rate of loan interest rate which is excluded in column (1) and included as a continuous variable in column (2) and as a discrete variable in column (3). Column (4) includes the interaction term between loan interest rate and monthly fixed effects.

§ The standard errors are clustered by the borrower's quality.

* Indicates statistically significant at the 0.1 level.

** Indicates statistically significant at the 0.5 level.

*** Indicates statistically significant at the 0.01 level.

Table 2.C.6. Default Equations without Control Variables: Least Squares Estimates

	(1)	(2)	(3)	(4)
VI: marital status	-0.0800***	-0.0800***	-0.0729***	-0.0611***
	(0.0232)	(0.0232)	(0.0215)	(0.0179)
VI: education	-0.0709***	-0.0709***	-0.0699***	-0.0702***
	(0.0109)	(0.0109)	(0.0104)	(0.0109)
VI: employment	0.136***	0.136***	0.121***	0.118***
	(0.0392)	(0.0392)	(0.0380)	(0.0403)
Married	-0.0232***	-0.0232***	-0.0221***	-0.0190**
	(0.00809)	(0.00809)	(0.00792)	(0.00771)
Associate degree	-0.0309**	-0.0309**	-0.0300**	-0.0339***
	(0.0125)	(0.0125)	(0.0126)	(0.0125)
Bachelor or higher	-0.124***	-0.124***	-0.119***	-0.121***
	(0.0255)	(0.0255)	(0.0254)	(0.0257)
Above median income($800)	0.0106	0.0106	0.0104	0.0156
	(0.0111)	(0.0111)	(0.0109)	(0.00980)
Agri	0.0661*	0.0661*	0.0689*	0.0806**
	(0.0381)	(0.0381)	(0.0394)	(0.0377)
Con, En, Ut	0.0387*	0.0387*	0.0425**	0.0320*
	(0.0230)	(0.0230)	(0.0216)	(0.0191)
Ret, whl, Tran	0.00267	0.00267	0.00739	0.00367
	(0.0140)	(0.0140)	(0.0126)	(0.0117)
Fin, law, RE, IT	-0.0542***	-0.0542***	-0.0479***	-0.0520***
	(0.0120)	(0.0120)	(0.0129)	(0.0122)
Edu, Health	-0.00897	-0.00897	-0.00275	-0.00869
	(0.0139)	(0.0139)	(0.0133)	(0.0128)
A&E, Sports, Hosp	0.0129	0.0129	0.0125	0.00395
	(0.0162)	(0.0162)	(0.0150)	(0.0138)
NGO, others	-0.0217	-0.0217	-0.0161	-0.0240*
	(0.0140)	(0.0140)	(0.0131)	(0.0126)
Gov	0.00259	0.00259	0.00953	0.00153
	(0.0163)	(0.0163)	(0.0150)	(0.0150)
Loan interest rate	No	Continous	Categorical	Interacted w/ Mon. FE
No control	✓	✓	✓	✓
R^2	0.077	0.077	0.093	0.126
Obs.	8,547	8,547	8,547	8,547

† Loan interest rate is the annualized percentage rate of loan interest rate which is excluded in column (1) and included as a continuous variable in column (2) and as a discrete variable in column (3). Column (4) includes the interaction term between loan interest rate and monthly fixed effects.

§ The standard errors are clustered by the borrower's quality.

* Indicates statistically significant at the 0.1 level.

** Indicates statistically significant at the 0.5 level.

*** Indicates statistically significant at the 0.01 level.

Table 2.C.7. Correlation among the Verification Indicators

	marital status	education	employment	house ownership	car ownership	ID	residence	certification	video-chat	social network	credit-report	cell-phone
marital status	1.00											
education	0.08	1.00										
employment	0.15	0.11	1.00									
house ownership	0.25	0.06	0.12	1.00								
car ownership	0.33	0.04	0.15	0.24	1.00							
ID	-0.01	-0.02	0.05	-0.02	-0.02	1.00						
residence	0.24	0.05	0.10	0.34	0.17	-0.03	1.00					
certification	0.07	0.10	0.07	0.07	0.03	-0.07	0.10	1.00				
video-chat	0.25	0.07	0.28	0.20	0.25	-0.02	0.15	0.05	1.00			
social network	0.13	0.17	0.08	0.06	0.06	-0.01	0.08	0.04	0.12	1.00		
credit-report	0.06	0.05	0.39	0.05	0.06	0.09	0.07	0.03	0.07	0.01	1.00	
cell-phone	0.25	0.14	0.10	0.14	0.16	-0.02	0.25	0.09	0.18	0.16	0.06	1.00

Table 2.C.8. Funding Equation: Least Squares Estimates of Control Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VI: ID	-0.00615	-0.00983	-0.0119	-0.00639	-0.00560	-0.00921	-0.0114	-0.00560
	(0.0213)	(0.0220)	(0.0220)	(0.0217)	(0.0212)	(0.0217)	(0.0218)	(0.0216)
VI: credit report	0.00768	0.00642	0.00644	0.00599	0.00717	0.00591	0.00593	0.00547
	(0.0160)	(0.0161)	(0.0162)	(0.0162)	(0.0156)	(0.0156)	(0.0157)	(0.0157)
VI: social network	-0.0112	-0.0127	-0.0110	-0.00873	-0.0109	-0.0125	-0.0108	-0.00856
	(0.0151)	(0.0147)	(0.0144)	(0.0143)	(0.0148)	(0.0144)	(0.0141)	(0.0141)
VI: cell-phone	-0.00490	-0.00384	-0.00325	-0.00285	-0.00507	-0.00396	-0.00336	-0.00292
	(0.00813)	(0.00773)	(0.00773)	(0.00795)	(0.00799)	(0.00758)	(0.00760)	(0.00780)
VI: video-chat	0.146***	0.132***	0.133***	0.129***	0.149***	0.135***	0.135***	0.131***
	(0.0132)	(0.0133)	(0.0129)	(0.0128)	(0.0131)	(0.0131)	(0.0127)	(0.0126)
VI: prof. certificate	-0.0348**	-0.0369***	-0.0360**	-0.0329**	-0.0342**	-0.0364**	-0.0354**	-0.0322**
	(0.0141)	(0.0141)	(0.0141)	(0.0142)	(0.0143)	(0.0143)	(0.0143)	(0.0143)
VI: car ownership	0.00418	0.00108	-0.00145	-0.00254	0.00591	0.00286	0.000335	-0.000704
	(0.0128)	(0.0124)	(0.0119)	(0.0121)	(0.0121)	(0.0117)	(0.0112)	(0.0113)
VI: house ownership	0.00782	0.00690	0.00661	0.00581	0.00853	0.00759	0.00731	0.00651
	(0.0129)	(0.0127)	(0.0124)	(0.0130)	(0.0129)	(0.0127)	(0.0124)	(0.0130)
VI: residence	-0.0126*	-0.00858	-0.00745	-0.00604	-0.0126*	-0.00855	-0.00740	-0.00596
	(0.00723)	(0.00724)	(0.00731)	(0.00711)	(0.00717)	(0.00716)	(0.00723)	(0.00702)
Car ownership	0.0199**	0.0180*	0.0179*	0.0163*	0.0206**	0.0185*	0.0184**	0.0169*
	(0.00976)	(0.00967)	(0.00932)	(0.00959)	(0.00959)	(0.00949)	(0.00913)	(0.00941)
Auto loan	-0.000958	0.00264	0.00274	0.00714	-0.00215	0.00149	0.00159	0.00597
	(0.00758)	(0.00745)	(0.00726)	(0.00750)	(0.00764)	(0.00751)	(0.00732)	(0.00752)
House ownership	0.000397	0.000251	-8.81e-05	-0.000719	-0.00170	0.0101	0.0118	0.0118
	(0.00622)	(0.00630)	(0.00637)	(0.00654)	0.000148	-7.69e-06	-0.000349	-0.000998
Mortgage	0.0143**	0.0113*	0.00999	0.0127**	0.0152**	0.0122*	0.0109*	0.0136**
	(0.00655)	(0.00676)	(0.00630)	(0.00586)	(0.00653)	(0.00673)	(0.00626)	(0.00582)

Continued on next page

Table 2.C.8. Funding Equation: Least Squares Estimates of Control Variables – Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Work experience:								
<1 yr	0.0665	0.0647	0.0546	0.0522	0.0596	0.0581	0.0480	0.0455
	(0.0695)	(0.0679)	(0.0681)	(0.0667)	(0.0681)	(0.0666)	(0.0669)	(0.0648)
1-3 yr	0.0880	0.0864	0.0756	0.0732	0.0804	0.0790	0.0681	0.0657
	(0.0713)	(0.0694)	(0.0695)	(0.0684)	(0.0696)	(0.0678)	(0.0680)	(0.0662)
3-5 yr	0.113	0.111	0.101	0.0981	0.105	0.103	0.0932	0.0900
	(0.0728)	(0.0709)	(0.0710)	(0.0701)	(0.0709)	(0.0691)	(0.0693)	(0.0677)
>5 yr	0.139*	0.135*	0.124*	0.120*	0.131*	0.128*	0.117*	0.112
	(0.0746)	(0.0721)	(0.0722)	(0.0713)	(0.0726)	(0.0702)	(0.0704)	(0.0688)
City size of residence:								
Small city	-0.0280	-0.0307	-0.0310	-0.0308	-0.0257	-0.0284	-0.0287	-0.0285
	(0.0337)	(0.0339)	(0.0339)	(0.0336)	(0.0326)	(0.0328)	(0.0327)	(0.0325)
Mid-sized city	-0.00411	-0.00709	-0.00540	-0.00267	-0.00657	-0.00944	-0.00786	-0.00538
	(0.0178)	(0.0177)	(0.0179)	(0.0159)	(0.0182)	(0.0182)	(0.0184)	(0.0162)
Metropolitan	-0.00963**	-0.00998**	-0.00993**	-0.00960**	-0.00907**	-0.00946**	-0.00941**	-0.00908**
	(0.00457)	(0.00450)	(0.00417)	(0.00427)	(0.00460)	(0.00452)	(0.00420)	(0.00429)
Location of residence:								
Northeast China	0.00812	0.00889	-0.00295	0.00536	0.00714	0.00803	-0.00390	0.00443
	(0.0606)	(0.0619)	(0.0577)	(0.0564)	(0.0585)	(0.0598)	(0.0556)	(0.0543)
East China	-0.00233	-0.00240	-0.0122	-0.00379	-0.00325	-0.00317	-0.0130	-0.00456
	(0.0499)	(0.0517)	(0.0481)	(0.0462)	(0.0481)	(0.0498)	(0.0464)	(0.0443)
Central China	-0.0188	-0.0195	-0.0286	-0.0200	-0.0193	-0.0199	-0.0290	-0.0203
	(0.0500)	(0.0516)	(0.0484)	(0.0463)	(0.0483)	(0.0498)	(0.0468)	(0.0446)
Western China	-0.00871	-0.00866	-0.0171	-0.00846	-0.00919	-0.00905	-0.0176	-0.00882
	(0.0498)	(0.0517)	(0.0483)	(0.0460)	(0.0481)	(0.0498)	(0.0466)	(0.0442)
Firm size (employees):								

Continued on next page

Table 2.C.8. Funding Equation: Least Squares Estimates of Control Variables – Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<10	-0.0461	-0.0607	-0.0587	-0.0619	-0.0234	-0.0380	-0.0360	-0.0383
	(0.0741)	(0.0680)	(0.0670)	(0.0646)	(0.0698)	(0.0641)	(0.0632)	(0.0609)
10-100	-0.0216	-0.0357	-0.0346	-0.0392	0.00114	-0.0131	-0.0120	-0.0157
	(0.0752)	(0.0690)	(0.0677)	(0.0651)	(0.0719)	(0.0662)	(0.0649)	(0.0626)
100-500	-0.00934	-0.0224	-0.0217	-0.0247	0.0133	8.39e-05	0.000709	-0.00130
	(0.0725)	(0.0663)	(0.0655)	(0.0628)	(0.0686)	(0.0627)	(0.0620)	(0.0595)
>500	-0.00665	-0.0212	-0.0208	-0.0229	0.0174	0.00265	0.00307	0.00201
	(0.0713)	(0.0653)	(0.0645)	(0.0621)	(0.0667)	(0.0611)	(0.0603)	(0.0583)
Loan purpose stated in description:								
Decoration	-0.0251	-0.0240	-0.0322	-0.0353	-0.0267	-0.0257	-0.0339	-0.0370
	(0.0780)	(0.0788)	(0.0765)	(0.0766)	(0.0794)	(0.0803)	(0.0780)	(0.0781)
Education/training	-0.0501	-0.0502	-0.0591	-0.0649	-0.0524	-0.0525	-0.0615	-0.0674
	(0.0938)	(0.0951)	(0.0928)	(0.0932)	(0.0957)	(0.0970)	(0.0948)	(0.0952)
Medical	-0.0758	-0.0671	-0.0746	-0.0735	-0.0771	-0.0684	-0.0759	-0.0747
	(0.0754)	(0.0747)	(0.0730)	(0.0728)	(0.0770)	(0.0762)	(0.0746)	(0.0744)
Other expenditure	-0.0398	-0.0404	-0.0489	-0.0528	-0.0413	-0.0420	-0.0505	-0.0545
	(0.0855)	(0.0860)	(0.0838)	(0.0842)	(0.0872)	(0.0878)	(0.0856)	(0.0861)
Car purchase	-0.0287	-0.0289	-0.0368	-0.0402	-0.0576	-0.0514	-0.0596	-0.0611
	(0.0766)	(0.0772)	(0.0745)	(0.0747)	(0.0861)	(0.0871)	(0.0844)	(0.0842)
Home purchase	-0.00713	-0.00890	-0.0174	-0.0202	-0.0303	-0.0306	-0.0385	-0.0420
	(0.0751)	(0.0760)	(0.0731)	(0.0743)	(0.0782)	(0.0788)	(0.0762)	(0.0764)
Short-term liquidity	-0.0767	-0.0760	-0.0846	-0.0883	-0.00869	-0.0105	-0.0190	-0.0218
	(0.0881)	(0.0889)	(0.0863)	(0.0864)	(0.0769)	(0.0778)	(0.0749)	(0.0761)
Entrepreneurship	-0.0187	-0.0189	-0.0271	-0.0301	-0.0777	-0.0770	-0.0856	-0.0893
	(0.0808)	(0.0814)	(0.0788)	(0.0791)	(0.0897)	(0.0904)	(0.0879)	(0.0880)
Wedding planning	-0.0218	-0.0199	-0.0299	-0.0311	-0.0198	-0.0201	-0.0282	-0.0313

Continued on next page

Table 2.C.8. Funding Equation: Least Squares Estimates of Control Variables – Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(0.0774)	(0.0792)	(0.0765)	(0.0772)	(0.0824)	(0.0830)	(0.0804)	(0.0808)
City GDP per capita	-6.74e-09	-1.01e-08	-8.59e-09	-2.51e-09	-2.02e-09	-5.80e-09	-4.12e-09	1.85e-09
	(5.66e-08)	(5.65e-08)	(5.63e-08)	(5.70e-08)	(5.61e-08)	(5.59e-08)	(5.58e-08)	(5.66e-08)
Loan principal	-4.91e-06***	-4.94e-06***	-4.90e-06***	-4.95e-06***	-4.97e-06***	-5.00e-06***	-4.95e-06***	-5.01e-06***
	(1.28e-06)	(1.27e-06)	(1.28e-06)	(1.29e-06)	(1.29e-06)	(1.28e-06)	(1.29e-06)	(1.30e-06)
Loan term	-0.00535***	-0.00456***	-0.00400***	-0.00378***	-0.00536***	-0.00456***	-0.00402***	-0.00379***
	(0.000672)	(0.000506)	(0.000720)	(0.000730)	(0.000673)	(0.000509)	(0.000725)	(0.000734)
Description length	0.0250***	0.0289***	0.0281***	0.0282***	0.0254***	0.0294***	0.0286***	0.0286***
	(0.00936)	(0.00991)	(0.00960)	(0.00947)	(0.00947)	(0.0100)	(0.00971)	(0.00958)
Positive words	0.00695*	0.00773**	0.00832**	0.00836**	0.00679*	0.00757**	0.00816**	0.00819**
	(0.00394)	(0.00379)	(0.00381)	(0.00405)	(0.00392)	(0.00377)	(0.00379)	(0.00404)
Top 100 words	0.0243*	0.0269*	0.0199	0.0179	0.0238*	0.0264*	0.0194	0.0174
	(0.0133)	(0.0141)	(0.0136)	(0.0135)	(0.0134)	(0.0142)	(0.0136)	(0.0136)
Controls of interaction terms:								
VI: marital status × Married					-0.0663	-0.0610	-0.0639	-0.0692
					(0.0499)	(0.0493)	(0.0502)	(0.0498)
VI: education ×								
Associate degree					0.0133	0.00357	0.00296	0.00236
					(0.0239)	(0.0236)	(0.0236)	(0.0237)
Bachelor or higher					0.0308	0.0219	0.0219	0.0220
					(0.0227)	(0.0219)	(0.0217)	(0.0220)
VI: income × above median income ($800)					-0.0425***	-0.0429***	-0.0432***	-0.0445***
					(0.0128)	(0.0127)	(0.0126)	(0.0130)

Continued on next page

Table 2.C.8. Funding Equation: Least Squares Estimates of Control Variables – Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Loan interest rate	No	Continuous	Categorical	Interacted w/ Mon. FE	No	Continuous	Categorical	Interacted w/ Mon. FE
R^2	0.359	0.366	0.373	0.383	0.36	0.367	0.373	0.383
Obs.	44,765	44,765	44,765	44,765	44,765	44,765	44,765	44,765

§ The standard errors are clustered by the borrower's quality.

* Indicates statistically significant at the 0.1 level.

** Indicates statistically significant at the 0.5 level.

*** Indicates statistically significant at the 0.01 level.

Table 2.C.9. Default Equation: Least Squares Estimates of Control Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VI: ID	0.0959**	0.116**	0.113**	0.113**	0.101**	0.122**	0.119**	0.119**
	(0.0460)	(0.0507)	(0.0496)	(0.0505)	(0.0470)	(0.0527)	(0.0517)	(0.0524)
VI: credit report	-0.0629	-0.0643	-0.0647	-0.0682*	-0.0624	-0.0638	-0.0642	-0.0678*
	(0.0393)	(0.0398)	(0.0397)	(0.0399)	(0.0392)	(0.0397)	(0.0396)	(0.0397)
VI: social network	-0.0631***	-0.0591***	-0.0569***	-0.0569***	-0.0645***	-0.0604***	-0.0581***	-0.0579***
	(0.0113)	(0.0111)	(0.0110)	(0.0109)	(0.0111)	(0.0110)	(0.0108)	(0.0108)
VI: cell-phone	-0.0158	-0.0181	-0.0181	-0.0164	-0.0158	-0.0182	-0.0181	-0.0163
	(0.0136)	(0.0136)	(0.0132)	(0.0132)	(0.0136)	(0.0135)	(0.0132)	(0.0132)
VI: video-chat	0.00469	0.00333	0.00340	0.00419	0.00605	0.00464	0.00475	0.00572
	(0.0146)	(0.0145)	(0.0144)	(0.0142)	(0.0145)	(0.0145)	(0.0144)	(0.0142)
VI: prof. certificate	0.0103	0.0102	0.00895	0.0131	0.00736	0.00739	0.00606	0.0103
	(0.0152)	(0.0152)	(0.0155)	(0.0168)	(0.0152)	(0.0152)	(0.0155)	(0.0168)
VI: car ownership	-0.0339***	-0.0287**	-0.0261**	-0.0273**	-0.0323***	-0.0270**	-0.0244**	-0.0256**
	(0.0117)	(0.0119)	(0.0119)	(0.0122)	(0.0117)	(0.0119)	(0.0118)	(0.0121)
VI: house ownership	-0.00702	-0.00700	-0.00749	-0.00424	-0.00709	-0.00702	-0.00752	-0.00452
	(0.0141)	(0.0142)	(0.0142)	(0.0139)	(0.0143)	(0.0144)	(0.0144)	(0.0141)
VI: residence	0.0186	0.0168	0.0150	0.0141	0.0192	0.0174	0.0157	0.0149
	(0.0131)	(0.0132)	(0.0132)	(0.0131)	(0.0129)	(0.0131)	(0.0130)	(0.0129)
Car ownership	-0.0475***	-0.0479***	-0.0475***	-0.0457***	-0.0484***	-0.0487***	-0.0484***	-0.0467***
	(0.0112)	(0.0111)	(0.0110)	(0.0113)	(0.0114)	(0.0113)	(0.0111)	(0.0114)
Auto loan	-0.0103	-0.00937	-0.0113	-0.0165	-0.0104	-0.00943	-0.0112	-0.0163
	(0.0144)	(0.0143)	(0.0139)	(0.0141)	(0.0142)	(0.0141)	(0.0138)	(0.0139)
House ownership	0.00377	0.00546	0.00551	0.00884	0.00229	0.00406	0.00415	0.00750
	(0.00754)	(0.00760)	(0.00769)	(0.00805)	(0.00765)	(0.00771)	(0.00785)	(0.00831)
Mortgage	-0.0477**	-0.0461**	-0.0470**	-0.0508***	-0.0490**	-0.0473**	-0.0481**	-0.0517***
	(0.0192)	(0.0187)	(0.0182)	(0.0177)	(0.0198)	(0.0193)	(0.0188)	(0.0182)

Continued on next page

Table 2.C.9. Default Equation: Least Squares Estimates of Control Variables – Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<1 yr	-0.209	-0.190	-0.202	-0.0588	-0.176	-0.158	-0.170	-0.0549
	(0.158)	(0.156)	(0.161)	(0.172)	(0.143)	(0.143)	(0.147)	(0.176)
1-3 yr	-0.203	-0.183	-0.194	-0.0496	-0.168	-0.150	-0.162	-0.0444
	(0.157)	(0.155)	(0.160)	(0.173)	(0.142)	(0.142)	(0.147)	(0.177)
3-5 yr	-0.225	-0.204	-0.215	-0.0722	-0.191	-0.172	-0.184	-0.0679
	(0.156)	(0.154)	(0.159)	(0.173)	(0.142)	(0.142)	(0.146)	(0.177)
>5 yr	-0.240	-0.219	-0.229	-0.0873	-0.207	-0.187	-0.198	-0.0834
	(0.157)	(0.155)	(0.160)	(0.173)	(0.142)	(0.142)	(0.147)	(0.178)
Small city	0.0566	0.0615	0.0604	0.0611	0.0544	0.0595	0.0582	0.0583
	(0.0426)	(0.0423)	(0.0427)	(0.0469)	(0.0415)	(0.0411)	(0.0416)	(0.0458)
Mid-sized city	0.0108	0.0159	0.0180	0.0300	0.0147	0.0197	0.0217	0.0343
	(0.0339)	(0.0337)	(0.0337)	(0.0337)	(0.0334)	(0.0332)	(0.0333)	(0.0334)
Metropolitan	-0.00345	-0.00325	-0.00304	-0.00408	-0.00539	-0.00510	-0.00487	-0.00588
	(0.0156)	(0.0159)	(0.0163)	(0.0161)	(0.0160)	(0.0163)	(0.0166)	(0.0164)
Northeast China	-0.0916	-0.0999*	-0.0792	-0.0702	-0.0877	-0.0961*	-0.0751	-0.0662
	(0.0617)	(0.0573)	(0.0594)	(0.0620)	(0.0622)	(0.0579)	(0.0599)	(0.0625)
East China	-0.154**	-0.165***	-0.143**	-0.135**	-0.149**	-0.160***	-0.137**	-0.129**
	(0.0607)	(0.0561)	(0.0591)	(0.0630)	(0.0606)	(0.0560)	(0.0589)	(0.0629)
Central China	-0.145**	-0.155***	-0.133**	-0.124*	-0.139**	-0.149***	-0.127**	-0.118*
	(0.0618)	(0.0572)	(0.0605)	(0.0639)	(0.0616)	(0.0570)	(0.0603)	(0.0637)
Western China	-0.183***	-0.195***	-0.174***	-0.168***	-0.179***	-0.191***	-0.169***	-0.164***
	(0.0612)	(0.0564)	(0.0593)	(0.0626)	(0.0611)	(0.0563)	(0.0590)	(0.0623)

Continued on next page

Table 2.C.9. Default Equation: Least Squares Estimates of Control Variables – Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<10	0.0351*	0.0320			0.0394*	0.0362*		
	(0.0198)	(0.0199)			(0.0212)	(0.0213)		
10-100	0.0450***	0.0445***	0.0125	0.00886	0.0473***	0.0467***	0.0106	0.00651
	(0.0113)	(0.0114)	(0.0161)	(0.0160)	(0.0116)	(0.0117)	(0.0166)	(0.0165)
100-500	0.0135	0.0123	-0.0195	-0.0222	0.0148	0.0136	-0.0223	-0.0253*
	(0.0102)	(0.0107)	(0.0145)	(0.0145)	(0.0102)	(0.0108)	(0.0154)	(0.0153)
>500			-0.0325	-0.0355*			-0.0367*	-0.0400*
			(0.0198)	(0.0203)			(0.0212)	(0.0216)
Decoration	-0.0548**	-0.0531*	-0.0550*	-0.0520	-0.0572**	-0.0553**	-0.0570**	-0.0540*
	(0.0269)	(0.0278)	(0.0288)	(0.0321)	(0.0263)	(0.0273)	(0.0284)	(0.0316)
Education/training	-0.154***	-0.156***	-0.159***	-0.164***	-0.160***	-0.161***	-0.164***	-0.169***
	(0.0248)	(0.0260)	(0.0258)	(0.0300)	(0.0247)	(0.0257)	(0.0255)	(0.0299)
Medical	-0.129***	-0.127***	-0.127***	-0.125***	-0.127***	-0.126***	-0.126***	-0.124***
	(0.0454)	(0.0424)	(0.0458)	(0.0455)	(0.0474)	(0.0446)	(0.0479)	(0.0476)
Other expenditure	-0.109***	-0.107***	-0.107***	-0.108***	-0.112***	-0.110***	-0.109***	-0.110***
	(0.0236)	(0.0245)	(0.0242)	(0.0280)	(0.0235)	(0.0243)	(0.0241)	(0.0280)
Car purchase	-0.0877***	-0.0901***	-0.0912***	-0.0887***	-0.0910***	-0.0931***	-0.0940***	-0.0912***
	(0.0254)	(0.0265)	(0.0270)	(0.0311)	(0.0248)	(0.0260)	(0.0265)	(0.0307)
Personal expenditure	-0.0476	-0.0449	-0.0469	-0.0458	-0.0517	-0.0487	-0.0506	-0.0495
	(0.0329)	(0.0334)	(0.0337)	(0.0367)	(0.0323)	(0.0329)	(0.0334)	(0.0363)
Home purchase	-0.0707*	-0.0681	-0.0704	-0.0662	-0.0760*	-0.0730*	-0.0752*	-0.0712
	(0.0420)	(0.0427)	(0.0441)	(0.0468)	(0.0415)	(0.0424)	(0.0439)	(0.0466)
Short-term liquidity	-0.0576**	-0.0537**	-0.0537**	-0.0512*	-0.0615***	-0.0572**	-0.0571**	-0.0546**
	(0.0232)	(0.0241)	(0.0240)	(0.0281)	(0.0226)	(0.0235)	(0.0236)	(0.0277)
Entrepreneurship	-0.0789***	-0.0765***	-0.0777***	-0.0758**	-0.0816***	-0.0789***	-0.0800***	-0.0779**

Continued on next page

Table 2.C.9. Default Equation: Least Squares Estimates of Control Variables – Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(0.0259)	(0.0269)	(0.0272)	(0.0311)	(0.0256)	(0.0267)	(0.0271)	(0.0311)
Wedding planning	-0.113***	-0.106***	-0.109***	-0.110***	-0.115***	-0.108***	-0.110***	-0.112***
	(0.0279)	(0.0276)	(0.0271)	(0.0272)	(0.0287)	(0.0283)	(0.0279)	(0.0282)
City GDP per capita	-3.71e-07***	-3.51e-07***	-3.46e-07***	-3.28e-07***	-3.84e-07***	-3.63e-07***	-3.58e-07***	-3.38e-07***
	(1.28e-07)	(1.31e-07)	(1.34e-07)	(1.25e-07)	(1.23e-07)	(1.26e-07)	(1.29e-07)	(1.20e-07)
Loan principal	-1.81e-06	-1.86e-06	-1.73e-06	-1.74e-06	-1.93e-06	-1.98e-06	-1.86e-06	-1.88e-06
	(1.48e-06)	(1.50e-06)	(1.52e-06)	(1.57e-06)	(1.55e-06)	(1.57e-06)	(1.59e-06)	(1.65e-06)
Loan term	0.00374***	0.00244***	0.00326***	0.00289***	0.00375***	0.00242***	0.00321***	0.00286***
	(0.000600)	(0.000688)	(0.000734)	(0.000803)	(0.000588)	(0.000682)	(0.000716)	(0.000784)
Description length	0.0295***	0.0273***	0.0276***	0.0263***	0.0289***	0.0267***	0.0269***	0.0256***
	(0.00676)	(0.00658)	(0.00666)	(0.00664)	(0.00682)	(0.00663)	(0.00672)	(0.00668)
Positive words	-0.00847	-0.00908	-0.00905	-0.00811	-0.00825	-0.00889	-0.00887	-0.00797
	(0.00605)	(0.00603)	(0.00595)	(0.00631)	(0.00602)	(0.00598)	(0.00592)	(0.00628)
Top 100 words	-0.00404	-0.00948	-0.0112	-0.0101	-0.00533	-0.0110	-0.0129	-0.0122
	(0.0259)	(0.0257)	(0.0260)	(0.0272)	(0.0255)	(0.0254)	(0.0255)	(0.0265)
Controls of interaction terms:								
VI: marital					0.0279	0.0247	0.0226	0.0301
status ×								
Married					(0.0548)	(0.0551)	(0.0551)	(0.0609)
VI: education ×								
Associate degree					-0.153***	-0.147***	-0.143***	-0.137***
					(0.0363)	(0.0372)	(0.0361)	(0.0366)
Bachelor or higher					-0.166***	-0.162***	-0.161***	-0.160***
					(0.0326)	(0.0325)	(0.0314)	(0.0320)
VI: income ×					-0.0103	-0.0115	-0.0126	-0.0144

Continued on next page

Table 2.C.9. Default Equation: Least Squares Estimates of Control Variables – Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
above median income ($800)					(0.0151)	(0.0151)	(0.0147)	(0.0151)
Loan interest rate	No	Continuous	Categorical	Interacted w/ Mon. FE	No	Continuous	Categorical	Interacted w/ Mon. FE
R^2	0.125	0.131	0.135	0.158	0.123	0.129	0.133	0.156
Obs.	8,547	8,547	8,547	8,547	8,547	8,547	8,547	8,547

§ The standard errors are clustered by the borrower's quality.

* Indicates statistically significant at the 0.1 level.

** Indicates statistically significant at the 0.5 level.

*** Indicates statistically significant at the 0.01 level.

Chapter 3

Credit Insurance Impacts on Peer-to-Peer Lending Markets: Evidence from China

3.1 Introduction

3.1.1 China: The World's Largest P2P Lending Market

Sprouting from two major economies, the United Kingdom and the United States, online lending businesses have expanded into many other advanced countries and emerging markets. Inspired by the business model developed in the UK and the US, PPdai–the Chinese version of Prosper– went live online in 2007, followed by the emergence of a number of marketplaces, including Hongling Captial, Lufax, and Renrendai. Figure 3.4.1 shows the exponential expansion of the Chinese P2P lending market in the last 3 years. The number of lending platforms has risen from about 100 to over 2,000, at a pace of 60 more platforms per month. The loan volume issued per month has grown from a round 10 billion yuan ($1.6 billion) at the beginning of 2014 to over 80 billion yuan ($12.8 billion) in July 2015, despite the decline of the average annual interest rate from above 20% to less than 14%. In 2014, the entire Chinese P2P market originated new loans worth over 250 billion yuan ($40 billion), and served over 630,000 borrowers and 1.16 million lenders, ranking number one in the global market.[1]

Figure 3.4.2 displays the geographic distribution of the registered offices of existing platforms by the end of September 2015. Similar to the economic geography of China, P2P platforms cluster within provinces that are wealthier and more developed. All of the three provinces (Guangdong, Shandong, and Zhejiang) and the two direct-controlled municipalities (Beijing and Shanghai) that have more than 200 lending platforms are located in coastal regions. On average each province has about 66 platforms registered locally, while most of the inland provinces have less than 50.

Credit rationing is acute in China, a transition economy with underdeveloped legal and financial systems (in the form of market segmentation, entry barriers, interest rate regulation, and capital account control). Researchers have found that, in contrast to state-owned and foreign-owned enterprises, Chinese private firms are subject to credit constraints that are significantly tighter despite the fact that the private sector constitutes the biggest engine of economic growth and the major provider of employment. This is believed to be the consequence of the discrimination of state-owned banks that dominate the traditional financial market as well as the lack of collateralizable assets (see Gregory and Tenev, 2001; Brandt and Li, 2003; Cull and Xu, 2003; Lu and Yao, 2009; Poncet, Steingress and Vandenbussche, 2010; Song, Storesletten and Zilibotti, 2011 and many others). Most bank loans in China require pledges of collateral, and the types of collateral acceptable to banks (and other traditional financial institutions) are mainly land and real estate that are not affordable for many individuals and SMEs. On the other hand, state-owned firms can acquire bank loans easily due to their political connections and privileges (see Bailey, Huang and Yang, 2011), while foreign-owned companies usually face loose credit constraints thanks to the government's investment-attraction policies and access to credit through internal financial market shared with their parent companies outside the border (see Manova, Wei and Zhang, 2015). According to *The World Bank Investment Climate Survey*, SMEs in China obtain only 12% of their capital from bank loans when their peers in Malaysia and Indonesia are able to obtain 21% and 24%, respectively. The credit constraint for Chinese SMEs deteriorates as firm size decreases. Bai et al. (2009) document that companies with at least 100 employees fund 27% of their capital from bank loans in China in comparison to 29%

[1] Data source: www.wangdaizhijia.com. Wangdaizhijia a third-party consulting firm that focuses on the Chinese P2P lending market.

in India, while the loan-to-capital share drops to 13% when the firm size is between 20 to 100 employees and to 2.3% when the firm size is 20 or fewer employees, compared to 38% and 29% in India. Ayyagari, Demirguc-Kunt and Maksimovic (2010) surveyed about 2,400 Chinese private firms only to find approximately 20% of them being able to finance their debt from banks.

Most research mentioned that the credit market allocation in China is far from what it would be in the efficient equilibrium, and this may result in a great demand for alternative (informal) financing means. Typically, informal financing lends small, unsecured, short-term loans to the demand from underdeveloped regions, less-profitable sectors, small entrepreneurial ventures, households and individuals, and its funding decision is inclined to rely more on personal relationships and reputation instead of formal collaterals. A lot of informal financial institutions are found to be more efficient than the traditional ones in monitoring their debtors and enforcing repayments because they have specialized in working with a particular group of borrowers. Ayyagari, Demirguc-Kunt and Maksimovic (2010) find that the informal financial institutions (e.g., family members, moneylenders, landlords) in China play an important role as a complement to the formal ones by serving the low end of the market. Trade credits, reputation and relationships are found to be the alternative financing channels that effectively support the growth of the private sector in China (see Allen, Qian and Qian, 2005; Ge and Qiu, 2007 and many others). Without well-established legal protection and with limited access to capital market, private firms in China still constitute the fastest growing sector due to their reliance on alternative financing and governance mechanisms as pointed in Linton (2006). Now with the surge of Internet, P2P lending is poised to be the next important channel of informal financing.

Aside from the demand of rationed individuals and firms, the development of P2P lending in China benefits from the desire for more investment opportunities, an important and probably unique feature of the Chinese market. Lardy (2008) demonstrates that the financial repression in China leads to low and even negative real interest rates, and hence depresses the growth of household income and causes serious loss in efficiency. Regardless, the debate on whether repressive financial policies are beneficial to economic growth and financial stability (see McKinnon, 1973; Stiglitz, 1994; Hellmann, Murdock and Stiglitz, 2000; Huang and Wang, 2011 and many others), financial repression lowers the opportunity cost of investment for both institutional and individual investors, which helps (at least partially) in explaining the rapid expansion of P2P lending in China.

3.1.2 Asymmetric Information and Credit Risk

Despite astonishing growth predictions, a number of significant issues present serious challenges to the development of the industry, among which fraudulent activities constitute the biggest challenge that can potentially stunt the growth of the business model. Alternative finance such as P2P lending is subject to severe asymmetric information problem and hence high credit risk. Most P2P loans are unsecured personal loans without any pledge of collaterals. While majority of P2P borrowers are individuals or small businesses without long and persistent records of good credit history. Even in the advanced economies with mature financial and information systems, a big part of borrower attributes are hard to be verified, many times even unverifiable, by online P2P platforms or investors. Background checking and information verification are in general much more limited for P2P lending compared to traditional intermediaries. For instance, when listing a loan request, a borrower can post most of her demographic information without any source of approval during the process (see Iyer et al., 2009; Michels, 2012). It is often very difficult for platforms or lenders to identify credible borrowers from bad ones due to such information asymmetry. Moreover, once a loan request is successfully funded on a P2P lending platform, the entire fund is transferred to its borrower shortly afterwards. Then it becomes even harder for platforms, let alone individual lenders, to monitor borrowers' use of funds and secure subsequent repayments.

As a result, credit risk has always been a serious challenge to the business model of P2P lending. Because there is no pledge of collaterals at all, credit risk is high in the market. As reported in Renton (2012), during the period from 2006 through October 2008, Prosper (or commonly referred to as "Prosper 1.0") issued 28,936 loans, all of which have since matured. 18,480 of these loans have been fully paid off while 10,456 loans defaulted leading to a default rate of 36.1%; $46,671,123 of the $178,560,222 loaned-out has been wiped off resulting in a loss rate of 26.1%. Detailed default rates are unknown for many Chinese platforms. However, based on the data collected by Wangdaizhijia, 1129 out of 3151 (more than 35%) platforms that ever established have turned to be malfunctioning[2] by July 2015, which implies an even higher investment risk in the Chinese P2P lending market.

[2]A platform is counted as malfunctioning if it fits into one of the following categories: (1) the platform's business or website is closed or unresponsive, (2) the platform has filed for bankruptcy, (3) the platform has a severe liquidity or insolvency problem, and (4) the platform is charged with fraud or other illegal activity by its users or the authority.

The severe problem of asymmetric information leads to high default risk and hence low funding probability as lenders are just risk averse. This in turn deters entry of borrowers with good quality which would result in even higher default risk, eventually leaving only "lemons" in the market. Therefore, in order to survive and grow business, P2P platforms compete fiercely in soliciting credible borrowers and default risk control. To address the asymmetric information issue, many platforms commit to improving their criteria for loan application and encouraging borrowers to provide more information for verification. Some platforms have incorporated social networking service in their algorithm to match borrowers with investors, based on a theory that borrowers are less likely to default to lenders with whom they have affinities and social relationships. To identify the common relationships among a pool of thousands of users, such matching algorithm often looks into factors such as geographic location, educational and professional background, and affiliation to a social network. Lin, Prabhala and Viswanathan (2013) study the dataset of loan listings on Prosper 1.0 and find that the online friendships of borrowers, as signals of credit quality, increase the probability of successful funding, lower interest rates on funded loans, and are associated with lower *ex post* default rates.

Alternatively, some Chinese P2P platforms have adopted an online-to-offline (O2O) business model in which they cooperate with offline financial institutions to provide credit guarantees or insurance on P2P loans. An offline guarantor often assists online platforms and lenders in searching for high-quality borrowers and verification of borrower characteristics. In case of a default, a guarantor would assume the complete or partial loan obligation by compensating lenders for their losses in principals and accrued interests. In this chapter, we are interested in studying the effects of loan guarantees on the P2P lending market – the market size, funding probability, and lenders' behavior in particular. In most cases, loan guarantees are purchased by borrowers voluntarily in the hope of improving their funding probability. To make a causal inference without the concern of endogeneity issue, we explore the quasi-experimental feature of staggered introduction of loan guarantees on a leading Chinese P2P platform across cities in different regions, which allows us to identify the impacts on loan guarantee on the market using a difference-in-difference method.

There has previously been little research on the effects of loan insurance in credit markets. Governments of many countries are inclined to extend financial access to SMEs because they believe that development of SMEs are crucial to job creation and future economic growth. Consequently, governments have often been the guarantors of loans that financial institutions advance to SMEs. Riding and Haines (2001) conduct a cost-benefit analysis of Canadian loan guarantee programs, Small Business Loan Act (SBLA). They conclude that default rates on the portfolio of guaranteed loans are particularly sensitive to the level of the guarantee. If loan guarantee level increases from 85% of principal and accrued interest to 90% of it, the expected default rate would increase 50%. Meanwhile, Cowling (2010) finds evidence for the existence of credit rationing in the bank loan market in the UK and the inception of SFLGS, a loan guarantee scheme initiated by the UK government, helps alleviate capital constraints on small firms. Instead of focusing on government programs and bank loans, our research studies the impacts of loan guarantees provided by private firms on a market that channels funds directly from individuals to individuals.

As a financial derivative, loan guarantee shares characteristics similar, though not identical, to those of credit default swap (CDS).[3] Both can be utilized as insurance against potential defaults and reduce capital suppliers' credit risk exposure. In most cases, loan guarantees are purchased by borrowers to solicit trust of lenders while CDSs are bought by creditors to offset potential losses in case of default.[4] And CDSs are likely to be traded for corporate bonds and sovereign debts. The existing literature focuses on the impacts of CDS on credit risk allocation, corporate financing, and credit supply. Jarrow (2011) argues that the trading of CDSs is welfare increasing because it facilitates a more optimal allocation of risks in the economy. Bolton and Oehmke (2011) find the use of CDS strengthens creditors' bargaining power and helps reduce the incidence of strategic default. Ashcraft and Santos (2009) find little evidence that CDS lowers the cost of capital for an average firm though it leads to a small reduction in bond and loan spreads of safer and more transparent firms. While Saretto and Tookes (2013) find CDS helps relax firm's capital constraint, especially in non-price contract terms, as firms with traded CDS contracts on their debt are able to maintain higher leverage ratios and longer debt maturities, and such effect is greatest in the periods when credit supply constraints are most

[3] A CDS is a financial swap agreement that the seller of the CDS will compensate the buyer (usually the creditor of the reference loan) in the event of a loan default (by the debtor) or other credit event. In the event of default the buyer of the CDS receives compensation (usually the face value of the loan), and the seller of the CDS takes possession of the defaulted loan.

[4] In general, anyone can purchase a CDS, even buyers who do not hold the loan instrument and who have no direct insurable interest in the loan. In this case, they are called "naked" CDSs. This is usually not true in the case of loan guarantee.

binding. A similar conclusion echoes in Hirtle (2009) that examines derivatives use by banks and credit provision at the bank portfolio level. Her findings suggest that greater use of financial derivatives leads banks to increase credit provision. Our contribution to the literature is that we measure the responses of both credit demand and supply to the inception of loan guarantee provision in a market full of unsophisticated borrowers and lenders. And the quasi-experiment feature that we explore provides us with strong causal inference.

Thanks to a quasi-experiment in the form of staggered introduction of loan guarantees on a Chinese P2P lending marketplace, we are able to estimate the treatment effects of credit insurance on credit demand and supply with a difference-in-difference model. Our estimates suggest significant, strong, and persistent loan guarantee impacts. In summary, the introduction of loan guarantees increases P2P loan demand by at least 300% with an increase in the average funding probability by over 60%. In addition, the time needed for funding speeds up by 170 hours while the average investment amount per lender increases by over $60.

3.2 Credit Insurance in P2P Lending Markets

In this chapter, we use a unique dataset of loan listings scrapped from Renrendai to estimate the impact of loan guarantees on the P2P lending market. As introduced in the previous chapter, Renrendai is among the first and largest Chinese P2P marketplaces in terms of customer market size as well as total loan volume. Up until the second quarter of 2015, the platform has issued over 150,000 loans with a total amount exceeding $1.4 billion. Total number of loans made within a quarter has quickly increased from under 1,000 in the early 2012 to over 25,000 in 2015; the quarterly volume exceeded $250 million at the second quarter of 2015. In 2014, Renrendai ranked as the 10th largest Chinese P2P lending platforms in terms of total transaction volume when it originated 61,265 loans worth over $600 million, a 138% increase from the previous year. For the first half of 2015, Renrendai issued 52,932 loans worth $500 million, a 186% increase from the first two quarters in 2014.[5]

[5] Source: Wangdaizhijia, www.wangdaizhijia.com.

3.2.1 Principal Protection and Loan Guarantees

Growth of the platform has been challenged by information asymmetry and default risk as so do many other marketplaces in both domestic and foreign markets. Despite the rapid increase in total loan volume, average funding probability had been low (less than 10%). Therefore, the platform is incentivized to raise average borrower quality so as to attract more investors and boost up funding probability as Renrendai makes most of its revenue from management fees charged on borrowers whose loans are successfully funded.[6] As mentioned in the previous chapter, the platform provides borrower information verification service and strongly encourages borrowers to verify their information such as marital status, education, employment, income, and property ownership. Information verification status is highlighted on each loan listing webpage to indicate borrower's credibility. For borrowers with more verified information, the platform lowers their borrowing rates as well as their monthly management fees. And as illustrated in the previous chapter, borrowers with more verified information earn more trust from lenders and receive higher funding probability.

In addition to the measure of information verification, Renrendai has also adopted credit insurance to the marketplace as a means to control loan risk expecting to further expand its market share. The platform initiated in April 2011 the principal-protection program providing partial credit insurance against loan default. By default, all Renrendai borrowers are required to participate in this program. For each consummated loan, a one-time premium is charged on the borrower when receiving the funds. The principal-protection premium costs between 0 to 5% of the loan principal depending on the borrower's quality (with respect to the quantity and quality of borrower information that is verified by the platform). The collected premiums are put together to form a risk pool. In case of a default, the principal-protection program will pay back the remaining principal (no interest payments included) to the lenders using funds from the risk pool while taking over the lenders' claims to the unpaid debt.[7] It is worth mentioning that the principal protection is only funded by the risk pool and therefore constrained by the depth of the pool. The platform itself does not promise any legal binding obligation to lenders and the insurance against loan defaults is halted whenever the pool runs out of funds.

[6] The rate of the management fee ranges from 0 to 0.8% of loan principal depending on individual borrower's quality. Borrowers pay the management fees to the platform as they repay their loans in monthly installments.

[7] Initially, according to the principal-protection program policy, loans were categorized as defaulted after being overdue for 90 days (or more). Later in March 2012, the program policy was changed and the current definition of default includes any loan that has been overdue for 30 days (or more).

Provision of unlimited loan guarantees started in December 2012 as Renrendai began cooperating in partnership with Ucredit, an offline financial company. As the guarantor of insured Renrendai borrowers, Ucredit agrees to perform the complete obligations of a guaranteed borrower (repaying both loan principal and interests accrued) should the borrower fails to do so to the lenders. Considering it the ultimate strategy to deal with the default risk, Renrendai actively promotes and encourages potential borrowers to purchase the guarantees as they request for loans through the platform. Figure 3.4.4 displays a sample webpage of a loan listing with the unlimited guarantee on Renrendai. It is highlighted in the red circles (on the top and the bottom of the webpage) that the loan listing is fully guaranteed by Ucredit. The guarantee premium is between 1 to 6% conditional on the borrower's quality perceived by the offline guarantor. For borrowers who purchase the guarantees from Ucredit, the requirement of participating in the principal protection program is waived.

3.2.2 Staggered Introduction of Loan Guarantees

The adoption of the principal protection at the online marketplace was instantaneous as all the borrowers are required to participate without exceptions. However, the introduction of the unlimited loan guarantees took place gradually. Unlike the mandatory principal protection, the unlimited guarantee has always been optional for Renrendai borrowers. When purchasing the guarantees from Ucredit, borrowers are required to submit certain documents including at least (1) proofs of residence and (2) proofs of employment in the city of residence for the most recent 6 months – in such a way, Ucredit also assists the platform in borrower information verification, which is shared with the platform and posted on loan listing pages.[8]

Due to legal constraints, Ucredit can only provide loan guarantees to borrowers who live and work in the city where the offline guarantor has a local branch office under operation. A borrower cannot purchase the unlimited guarantee if no Ucredit local office operates in the city of the borrower's residence. Prior to Renrendai's introduction of the unlimited guarantees, Ucredit had already been running business in 7 Chinese cities. Since then, Ucredit has been gradually expending its business. By June 2015, Ucredit service is available in 60 Chinese cities. Table

[8]Proofs of residence include a person's registration record in the Hukou system or residence permits/proofs issued by the local government. If working as an employee, the borrower needs to provide proofs of employment with a monthly income over $320 (2,000 yuan). Or as a business owner, the borrower can provide proofs such as business license registered at the local Administration for Industry and Commerce (AIC).

3.4.1 lists the cities where Ucredit operates local offices in addition to the opening date of the first local office in the city and city-level aggregate economic data such as population and GDP per capita.

3.3 Identification Strategy and Empirical Results

3.3.1 Quasi-Experiment and Difference-in-Difference Methods

The research question is to evaluate the impacts of credit insurance on the P2P lending market. In particular, the goal is to estimate the loan guarantee effects on number of loan listings, funding probability, and lenders' bidding behavior. Despite the platform's active promotion, it is the borrower's decision whether or not to purchase the loan guarantee and therefore involves essentially a self-selection process which can result in an endogeneity problem. A simple comparison between guaranteed and unguaranteed loans can lead to biased results as borrowers who choose to purchase guarantees may not be comparable to those who choose not. It is intuitive to consider that guaranteed borrowers have better overall quality than unguaranteed borrowers as they need to pass the background check by the guarantor in addition to the regular one by the online platform.

Fortunately, the staggered introduction of loan guarantees provides us with an opportunity to explore exogenous variation in a quasi-experiment that is very similar to the one studied in Jensen (2007). As discussed in the previous section, the availability of loan guarantee service is defined by city boundaries. As the guarantor gradually expanded its business covering more and more regions, it generated differences in the timing of loan guarantee service availability to borrowers residing in different cities. Assuming that relocation cost is too large for borrowers to move simply for obtaining loan guarantees, we can estimate the loan guarantee impacts by exploring the timing differences. Namely, the impacts can be measured as the difference in variables of interest between observations from cities with guarantee service (i.e., the treatment group) and those from cities without guarantee service (i.e., the control group) using a difference-in-difference method.

Data Sample in Use

The raw dataset is scrapped directly from the website of Renrendai including all borrower characteristics and loan information available on loan listing webpages and bidding information of lenders. The variables in use are publicly available to any platform user. In the raw dataset, it contains the attributes of more than 254,000 loan requests and the associated borrowers. The majority of borrowers have only posted one loan listing while no more than 1,000 borrowers have records of multiple posts. The earliest observation dates back to October 2010 when the platform was established and the latest observations are by the end of June 2014.

Loan inform includes borrowing purpose, amount of principal, annualized interest rate, borrowing term (in months), and dummy indicators of its funding status and whether it is principal protected or unlimited guaranteed. Borrower attributes are represented by categorical variables including age, education level, marital status, monthly income, employment, city of residence, occupation, work experience, property ownership, car ownership, records of mortgage and auto-loans.

As implied in Table 3.4.1, cities that obtain guarantee service availability earlier tend to be larger and have higher GDP per capita. Thus, we trim down the raw dataset by excluding loan listings and borrowers from cities without available guarantee service by June 2014 – the end of our sample period. In addition, to focus on the short-run effects of loan guarantee, we restrict our dataset to include only observations that happened 30 weeks prior to and post the opening date of Ucredit local office. As a result, the number of total observations in the sample in use decreases to slightly under 40,000 containing either principal-protected or unlimitedly-guaranteed loan listings (hereafter referred as PP loans and UG loans) that were posted by borrowers from 31 cities. These cities in our sample can be divided into 12 groups, each group defined by the week of the Ucredit office open-date. Therefore, the time span is divided into 13 time periods (of different lengths). Within the first period, no cities had guarantee service available. Then as time progresses into the next period, borrowers from 1 to 4 additional cities would become eligible to purchase loan guarantees. At the end of the sample, borrowers from all 31 cities are capable of purchasing loan guarantees.

Empirical Model

Using a difference-in-difference method to estimate the treatment effects of loan guarantees on P2P lending market, we can specify the regression model as follows:

$$y_{igt} = \delta(t) + \sum_{w=-T}^{T} \beta_w UG_{igt}^{w} + FE_i + FE_g + \varepsilon_{igt} \tag{3.1}$$

where y_{igt} is an outcome variable of interest in city i of group g and at time t. UG_{igt}^{w} is a dummy variable indicating whether or not city i of group g has loan guarantee service available at time t; while the superscript w denotes the difference (in weeks) between time t and the week when guarantee service becomes available in city i.

FE_i and FE_g denote city and group fixed effects. In the following analysis, the time trend $\delta(t)$ is allowed to be either linear or stochastic. The linear time trend is captured by substituting $\delta(t) = \alpha_0 + \alpha_1 t$ into the regression model, i.e.,

$$y_{igt} = \alpha_0 + \alpha_1 t + \sum_{w=-T}^{T} \beta_w UG_{igt}^{w} + FE_i + FE_g + \varepsilon_{igt}. \tag{3.2a}$$

The stochastic time trend is instead measured by including weekly time fixed effects in the regression model, i.e.,

$$y_{igt} = \sum_{w=-T}^{T} \beta_w UG_{igt}^{w} + FE_i + FE_g + FE_t + \varepsilon_{igt}. \tag{3.2b}$$

Therefore, the coefficient β_w captures the w^{th} weekly treatment effect loan guarantee on the outcome variable y. In the empirical analysis, the time t is measured in weeks and we estimate weekly effects until the 30th week (more than 6 months) after loan guarantee becomes available, i.e., $T = 30$, such that we are able to measure not only the immediate effects but also the impacts in the medium run. Additionally, we also estimate the weekly effects during the 30 weeks before loan guarantee is available by including β_w for $w < 0$ in the regression model. This allows us to run a placebo test by comparing the pre-treatment sequences of the outcome variable between the control and treatment groups. If the difference-in-difference specification is valid, then we ought to have $\beta_w = 0$ for $w < 0$.

3.3.2 Empirical Results

Main Results

We use number of loan listings to represent the market demand for P2P lending. In particular, we estimate the regression Equation (3.1) with the outcome variable being the total number of loan listings posted by borrowers from city i of group g in week t. The estimation results are graphically presented in the upper panel of Figure 3.4.5 and recorded in column (1-2) in Table 3.4.3. It is apparent that the treatment effect of loan guarantee on the credit demand is significant and the magnitude is huge. In column (3-4) and (5-6) in Table 3.4.3, it also includes the regression results when outcome variable is changed to the logarithm of the total number of loan listings and share of UG loans among all loans. Once the loan guarantee becomes available, the number of loan posts increases by 20 to 30 per week per city or roughly 300% to 400% in percentage terms. As a result, UG loans account for an average 75% of the total loan posts after guarantee becomes available. The treatment effect reaches its full size within 3 weeks after treatment takes place and lasts persistently for 30 weeks without signs of decaying. The same estimation is conducted for number of PP loans only and the results are shown in the lower panel of Figure 3.4.5 and in columns (1-4) in Table 3.4.4. However, no significant changes are found. Therefore, the increase in the market demand for P2P lending consists of mostly new additional borrowers who are willing to purchase guarantees; while the introduction of loan guarantees has no effect on the borrowing behavior of borrowers who are not willing to purchase guarantees.

The supply in the P2P market is characterized by the average funding probability. In this case, Equation (3.1) is estimated using the fractional response regression method as the outcome variable is the average funding rate per loan posted by borrowers from city i of group g in week t.[9] In Figure 3.4.6 and Table 3.4.5, it exhibits the treatment effect of loan guarantee on funding probability. When loan guarantee becomes available, average funding probability increases significant and persistently by around 70%. The rise in the funding probability can only be attributed to the emergence of UG loans as no significant effect is found on the average funding probability of PP loans after guarantees become available. In effect, UG loans are so

[9]The fractional response regression is the proper way to estimate the model because the dependent variable represents essentially averages of a {0,1} binding variable. The method captures particular nonlinear relationships between the dependent and independent variables, especially when the outcome variable is near 0 or 1. The fractional response regression implements quasilikelihood estimators with the assumption of logit probability distribution.

popular among investors that their funding probability is nearly 100%; while that of PP loans remains constant (at around 5%) all the time. The same conclusion is drawn when we weight the funding probability by loan amount and use the average funding rate per dollar as the dependent variable in the regression instead. The results are shown in Figure 3.4.7 and Table 3.4.5.

Finally, we examine the treatment effect on lender's behavior by estimating the impacts on funding speed and average bidding amount. The funding speed measures the time needed for a loan list to be funded, which is defined as the time difference between the first and last bid (for funded loans). In Figure 3.4.8 and Table 3.4.6, it shows the strong and persistent treatment effect of loan guarantee on the city-week average funding speed: the average time required for funding a loan reduces significantly by 170 hours or equivalently a week of time as the result of the introduction of loan guarantees. While in Figure 3.4.9 and Table 3.4.6, it also displays the treatment effect of loan guarantee on the average bidding amount of lenders which is measured at the city-week averages for all funded loans. Under the treatment effect, the average bid amount increases by \$60 to \$70 showing a strong preference of lenders for UG loans over PP loans.

Results after Controlling for Borrower Information

Why are UG loans so much more favored by lenders relative to PP loans? Perhaps it is because borrowers of UG loans are much better in quality, i.e., the treatment effect is mostly a signaling effect. Looking carefully at the summary statistics of borrower attributes listed in Table 3.4.2, we dd find traces of differences in borrower quality. On average, borrowers of UG loans have relatively higher education level and month income than borrowers of PP loans, and they are more likely to own houses and have longer working experience. To find out how much of the loan guarantee treatment effect can be explained by the differences in borrower's quality, we regress Equation (3.1) again this time at the loan level instead of city average level. The dependent variable is the funding status of each loan, while all the observed borrower characteristics and loan information are also included as controls.[10] The estimation results are presented in Figure 3.4.10 and Table 3.4.7. Comparing the treatment effect estimates with and without controlling for the borrower characteristics and loan information, the treatment effect is slightly smaller when borrower's quality is controlled. However, the difference between the estimates is not

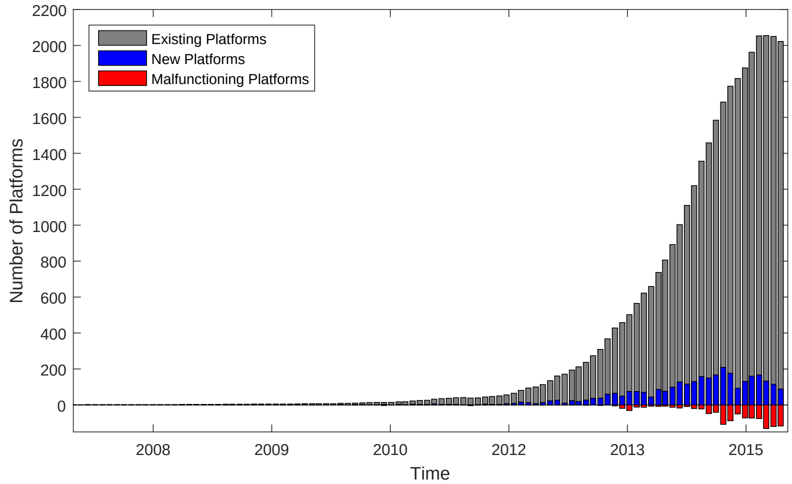
[10]Because the dependent variable is a dummy variable, the model is estimated as a logit model with MLE method.

comparable the over all magnitude of the effect. Therefore, the huge difference in the funding probability between UG and PP loans cannot be mostly explained by the difference in borrower's quality, i.e., the treatment effect is not merely a signaling effect.

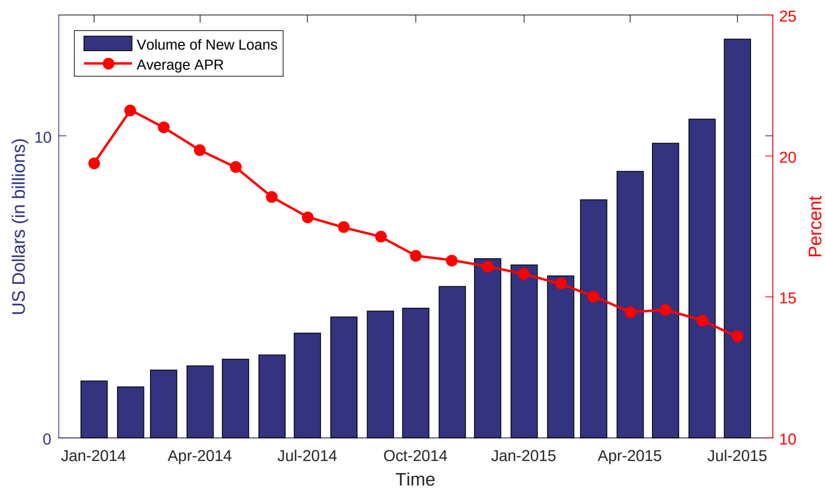
3.4 Conclusion

In this chapter, we estimate the treatment effects of credit insurance on demand and supply in P2P lending market. Thanks to a quasi-experiment on a Chinese P2P lending platform, we are able to identify the treatment effects with a difference-in-difference model by exploring the differences in the timing of loan guarantee service availability to borrowers residing in different cities. The estimates indicate signifiant, strong, and persistent loan guarantee treatment effects on P2P lending market, which result in a rise in the loan demand by over 300% with an increase in the average funding probability by over 60%. Meanwhile, there is a large drop in the time required for loans to be funded and big increase in lender's bidding amount. Although borrowers with loan guarantees are on average better in quality, this small change cannot fully explain the huge difference in the demand and supply. In future research, we can analyze the treatment effects on social welfare by incorporating data on guarantee contracts and lenders' and borrowers' outside options.

Figures and Tables



(a) Number of P2P Platforms

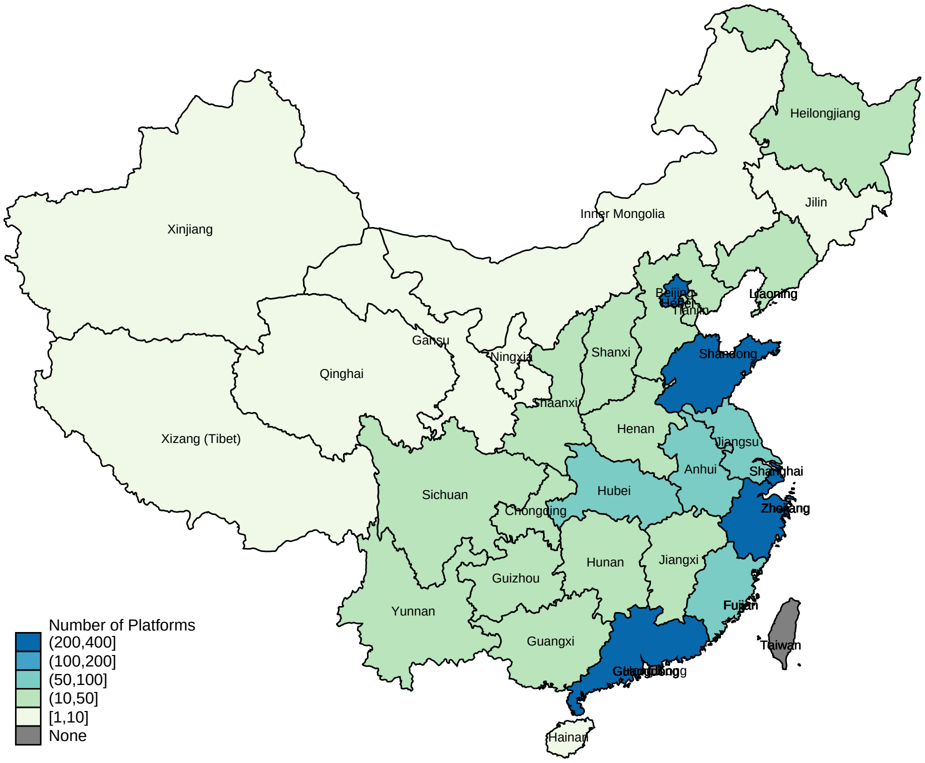


(b) Volume of New Loans and Average Loan Interest Rates

Note: (a) Data source: 01Caijing (available at http://data.01caijing.com/p2p/website/count.html). A platform is counted as malfunctioning if it fits into one of the following categories: (1) the platform's business or website is closed or unresponsive, (2) the platform has filed for bankruptcy, (3) the platform has a severe liquidity or insolvency problem, and (4) the platform is charged with fraud or other illegal activity by its users or the authority. The number of existing platforms is the cumulative sum of the new platforms net the problematic ones.

(b) Data source: Wangdaizhijia (available at http://shuju.wangdaizhijia.com/industry-list.html)

Figure 3.4.1. Development of P2P Lending Market in China



Note: Data source: 01Caijing (available at http://data.01caijing.com/p2p/website/map.html)

Figure 3.4.2. Geographic Distribution of P2P Lending Platforms in China

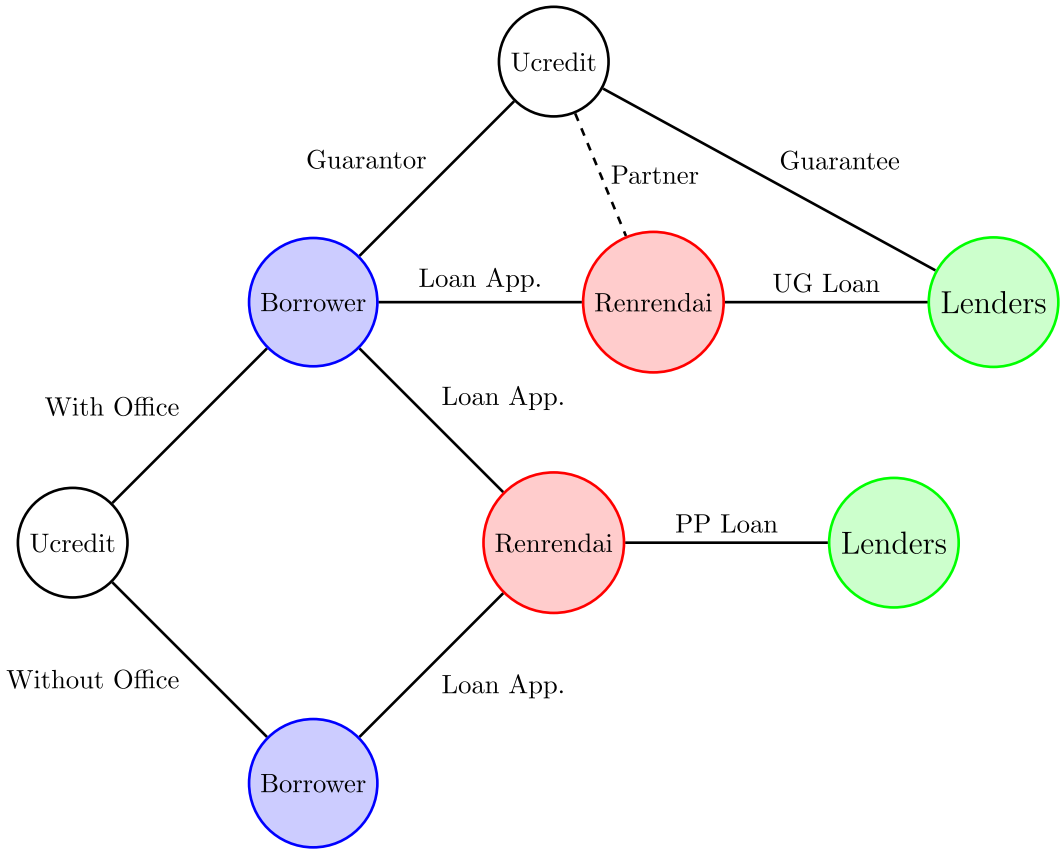
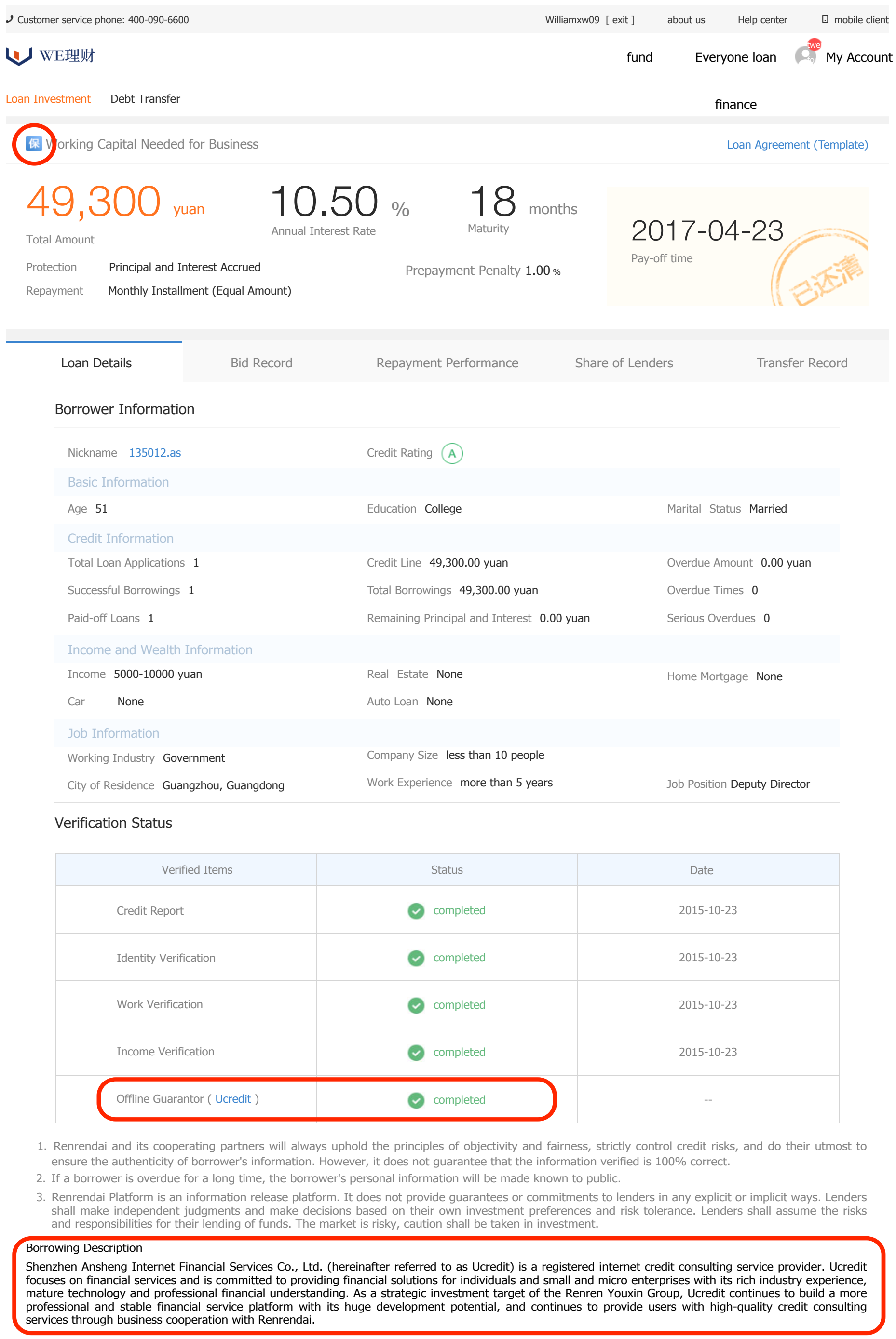


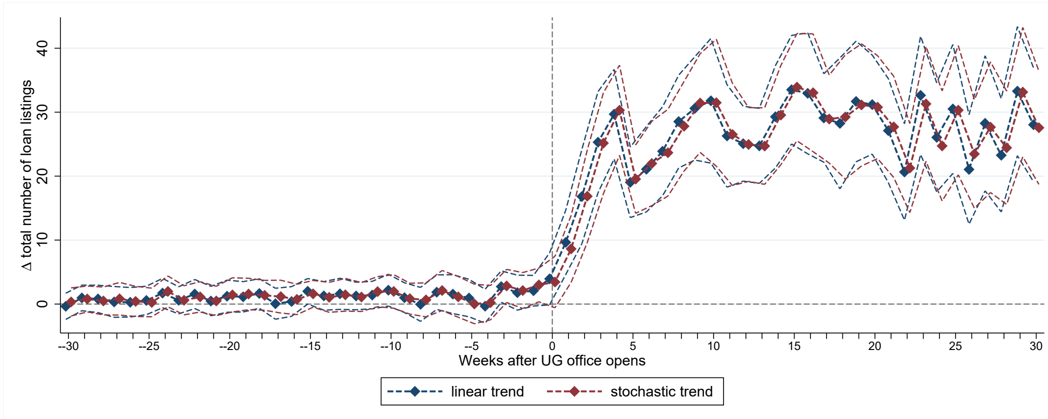
Figure 3.4.3. Borrowing Procedure at Renrendai



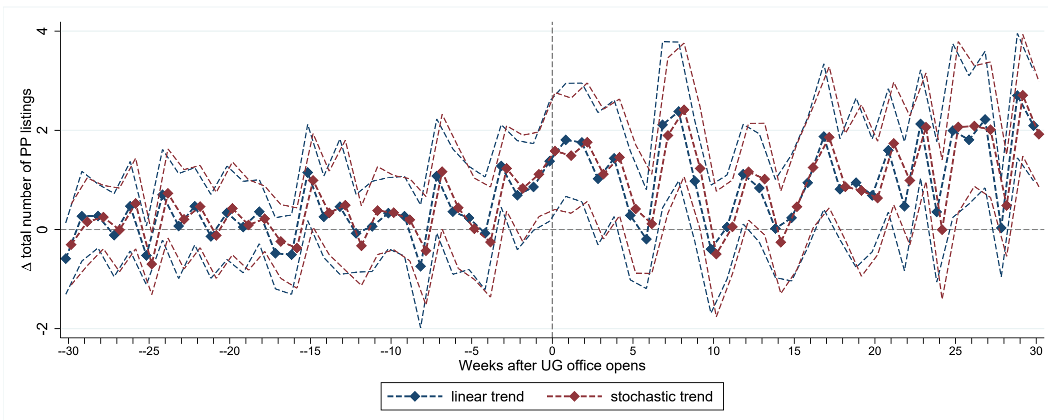
Note: (a) The sample has been translated from Chinese to English. The original page is available through http://www.renrendai.com/loan/184231.

(b) Within the red circles, it highlights that the loan listing is fully guaranteed by Ucredit.

Figure 3.4.4. Sample Webpage of a UG Loan Listing on Renrendai



(a) All loan listings

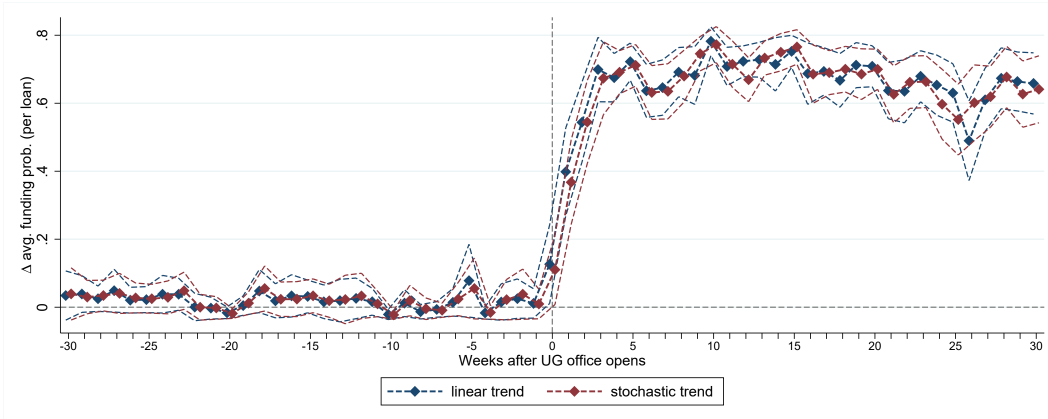


(b) PP loan listings

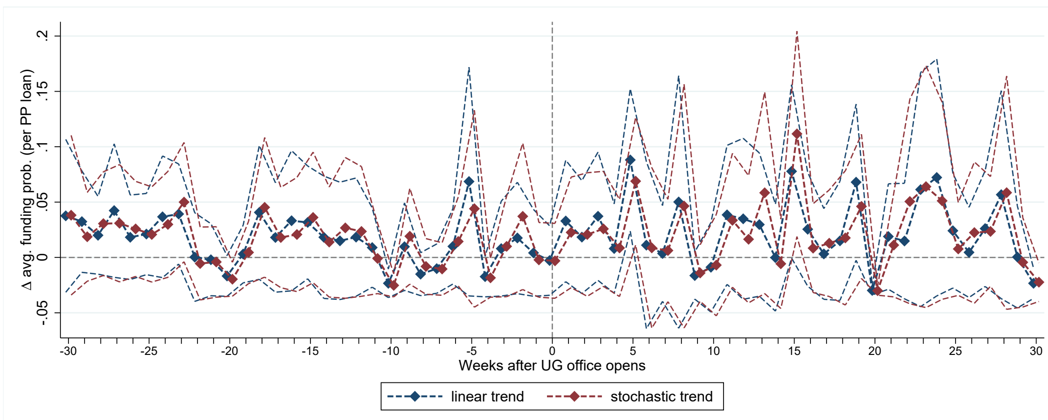
Note: § Numbers of loan listings are measured at city-week averages.

† Dashed lines are 95% confidence intervals.

Figure 3.4.5. Impact of Credit Insurance on Total Number of Loan Listings



(a) All listed loans

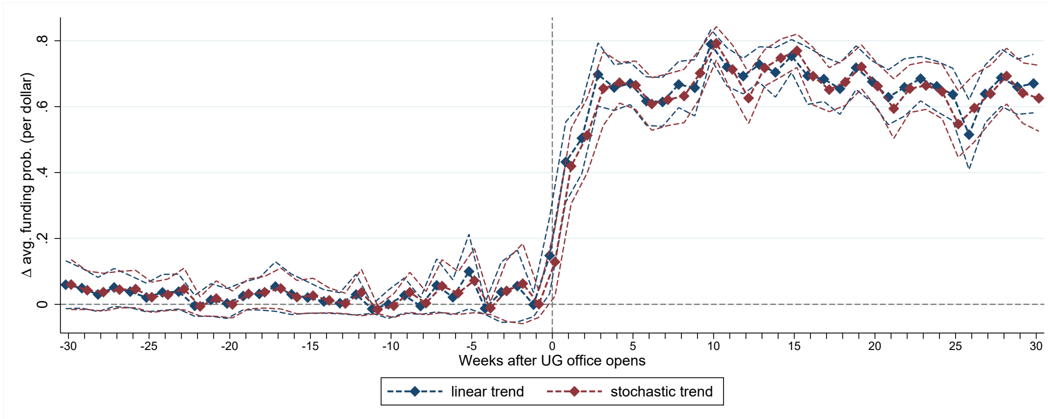


(b) Listed PP loans

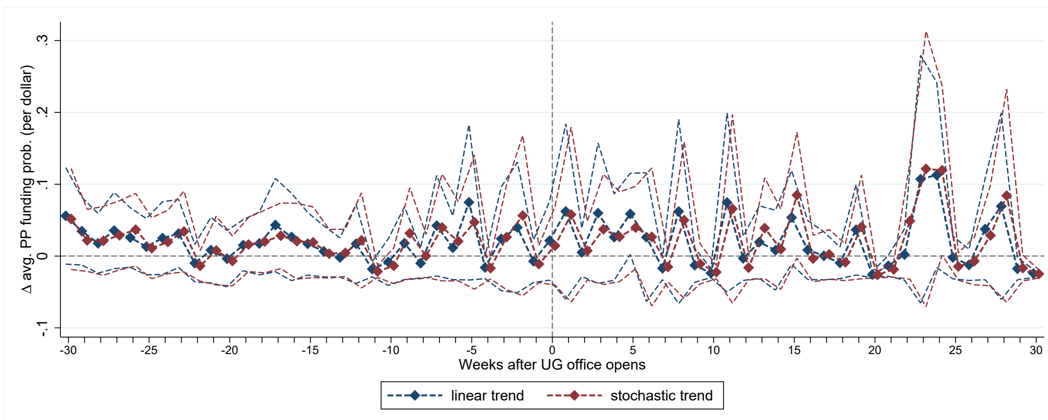
Notes: ∗ Funding probabilities are measured at city-week unweighted averages.

† Dashed lines are 95% confidence intervals.

Figure 3.4.6. Impact of Credit Insurance on Average Funding Probability (per loan)



(a) All listed loans

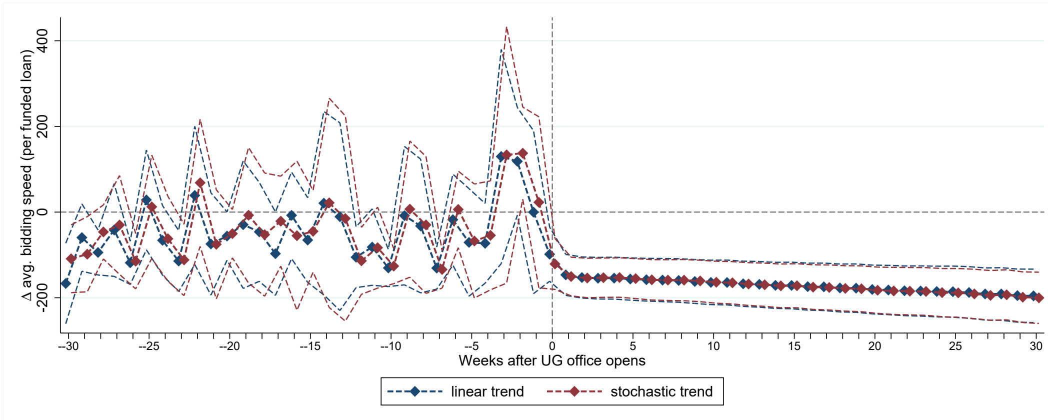


(b) Listed PP loans

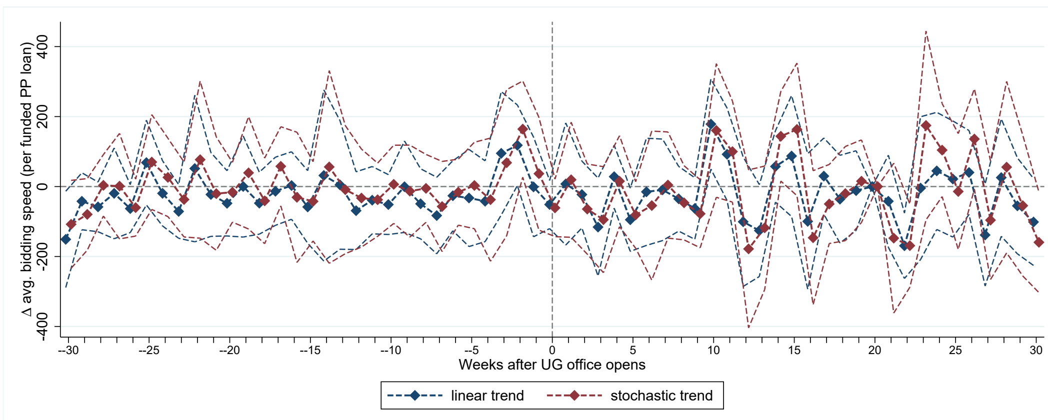
Notes: ∗ Funding probabilities are measured at city-week averages weighted by loan amount.

† Dashed lines are 95% confidence intervals.

Figure 3.4.7. Impact of Credit Insurance on Average Funding Probability (per dollar)



(a) All funded loans

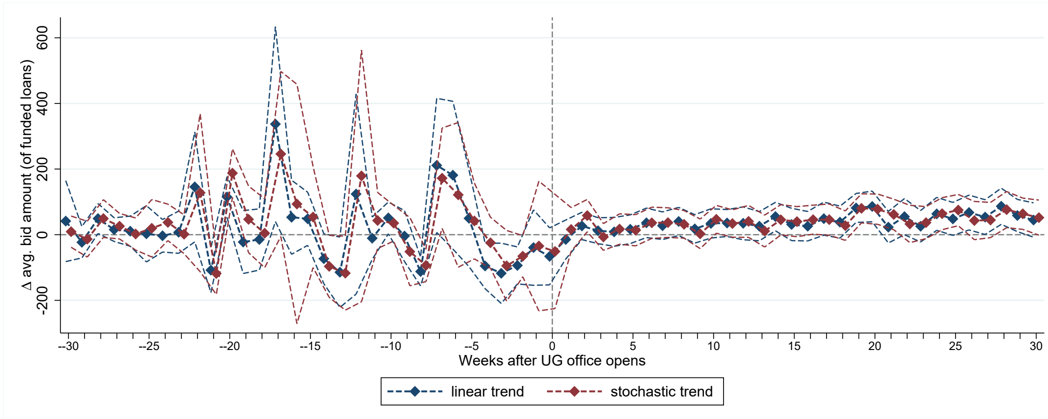


(b) Funded PP loans

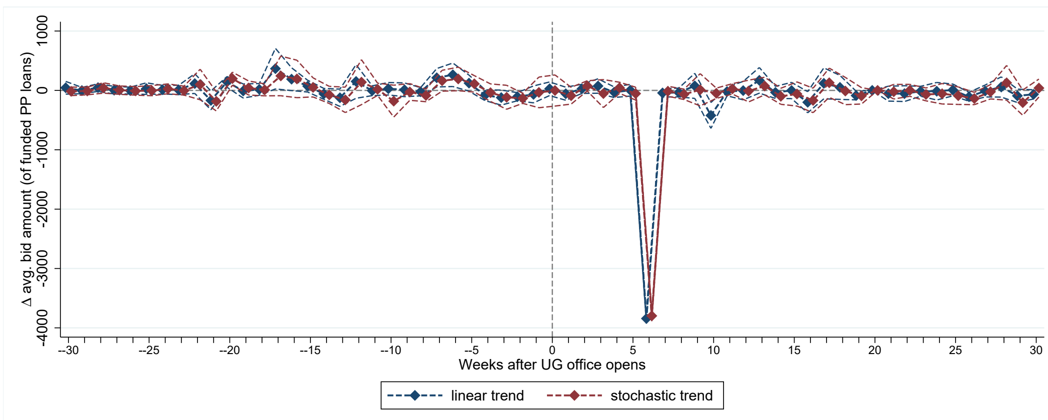
Notes: ∗ Funding speed is defined as the time difference between the first and the last bid and measured in hours at city-week unweighted averages.

† Dashed lines are 95% confidence intervals.

Figure 3.4.8. Impact of Credit Insurance on Average Funding Speed (per funded loan)



(a) All funded loans

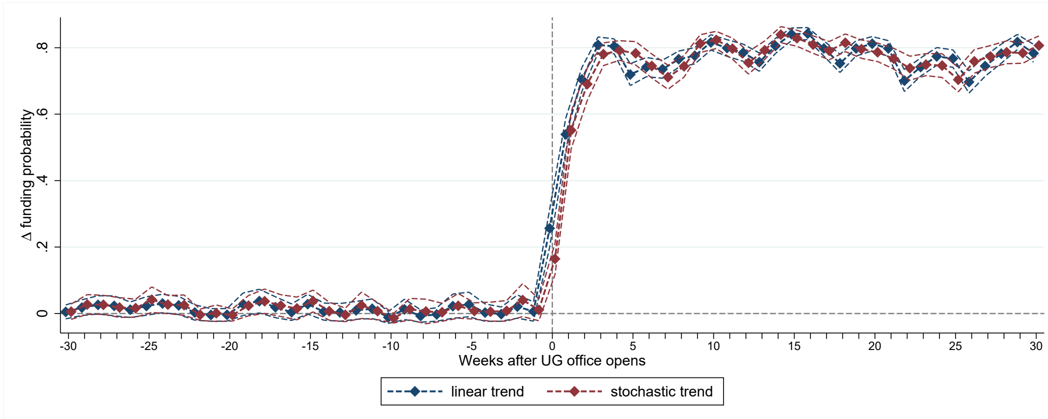


(b) Funded PP loans

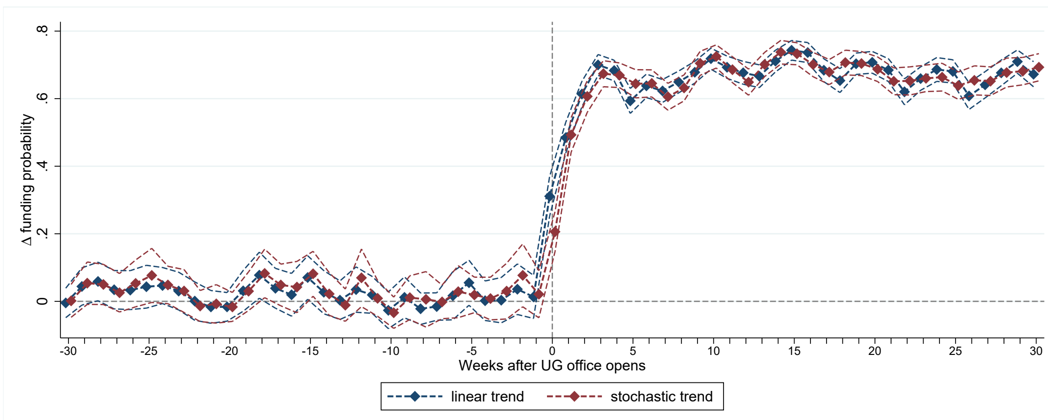
Notes: ∗ Bid amounts are measured in US dollars at city-week unweighted averages.

† Dashed lines are 95% confidence intervals.

Figure 3.4.9. Impact of Credit Insurance on Average Bid Amount (per funded loan)



(a) Without control for borrower information

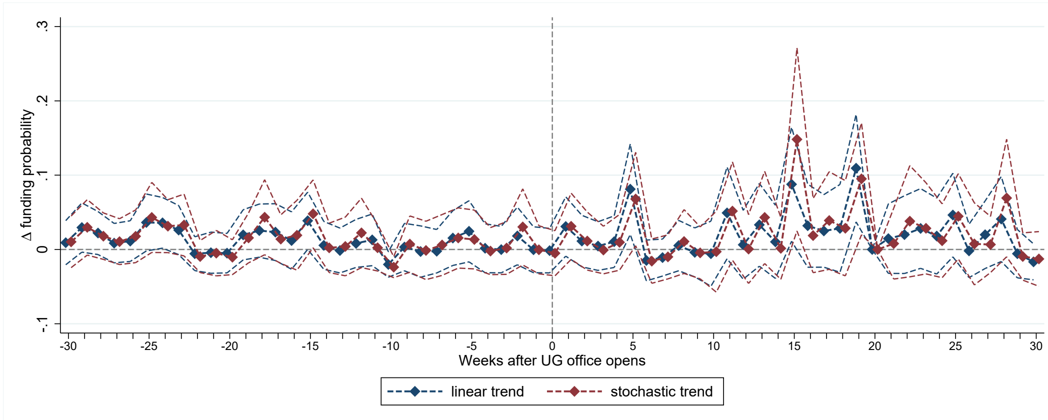


(b) With control for borrower information

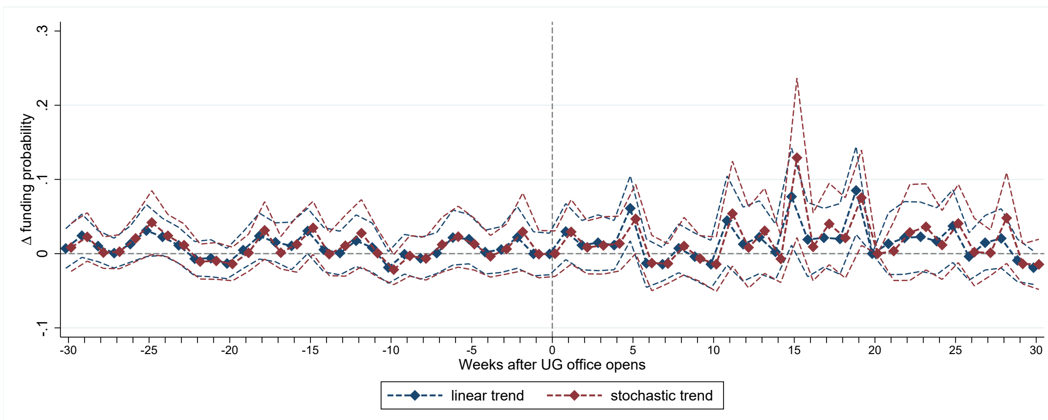
Notes: ∗ Funding probabilities are measured at city-week unweighted averages.

† Dashed lines are 95% confidence intervals.

Figure 3.4.10. Impact of Credit Insurance on Average Funding Probability (per loan)



(a) Without control for borrower information



(b) With control for borrower information

Notes: ∗ Funding probabilities are measured at city-week unweighted averages.

† Dashed lines are 95% confidence intervals.

Figure 3.4.11. Impact of Credit Insurance on Average Funding Probability (per PP loan)

Table 3.4.1. Waves of Ucredit Office Expansion

Wave	Date	City	Prov.	Pop.	GDP p.c.	Freq.	# of lists per week	Funding prob.	
								before	after
1	Nov 27, 2012	Beijing	BJ	2114.8	14914.08	4.57%	25.4	6.3%	7.8%
	Nov 27, 2012	Chengdu	SC	1429.8	10193.19	3.83%	21.4	2.5%	56.3%
	Nov 27, 2012	Chongqing	CQ	2970.0	6818.42	7.24%	40.3	2.1%	80.7%
	Nov 27, 2012	Nanjing	JS	818.8	15656.03	1.78%	10.1	4.2%	59.6%
	Nov 27, 2012	Shenyang	LN	825.7	13871.53	2.37%	13.9	2.9%	80.5%
	Nov 27, 2012	Suzhou	JS	1056.5	19712.00	2.70%	15.0	2.5%	50.4%
	Nov 27, 2012	Tianjin	TJ	1472.2	15617.51	2.91%	16.4	5.4%	74.1%
	Dec 1, 2012	Shanghai	SHG	2415.2	14311.07	8.97%	50.0	4.8%	68.8%
2	Jan 1, 2013	Guangzhou	GD	1292.7	19086.11	4.38%	24.4	5.0%	38.5%
	Jan 1, 2013	Qingdao	SD	896.4	14291.12	5.58%	31.4	4.4%	87.7%
	Jan 1, 2013	Xiamen	FJ	373.0	12946.52	2.35%	13.1	4.7%	66.4%
3	Mar 1, 2013	Wuhan	HUB	1022.0	14170.29	5.62%	31.3	2.5%	83.4%
	Mar 1, 2013	Zhengzhou	HEN	919.1	10796.47	4.79%	27.1	7.9%	83.2%
4	May 1, 2013	Changsha	HUN	722.1	15848.75	4.15%	23.5	2.1%	82.9%
	May 1, 2013	Dalian	LN	591.4	20698.78	3.62%	20.8	4.0%	90.9%
	May 1, 2013	Guiyang	GZ	452.2	7378.94	0.82%	5.0	2.5%	50.1%
	May 1, 2013	Kunming	YN	657.9	8305.97	0.78%	4.5	3.3%	23.9%
5	Jun 1, 2013	Fuzhou	FJ	734.0	10198.35	3.03%	16.9	3.4%	74.9%
	Jun 1, 2013	Yantai	SD	698.6	12857.28	2.63%	15.9	1.5%	86.8%
	Jun 1, 2013	Zunyi	GZ	614.3	4127.75	0.31%	2.8	2.6%	50.0%
6	Jul 1, 2013	Hefei	AH	761.1	9823.49	2.40%	13.8	3.1%	69.1%
	Jul 1, 2013	Xian	SAA	858.8	9099.35	3.42%	19.2	1.3%	82.4%
7	Aug 1, 2013	Lanzhou	GS	364.2	7804.40	1.24%	8.3	2.6%	87.3%
	Aug 1, 2013	Quanzhou	FJ	836.0	9986.61	2.72%	15.2	4.4%	63.9%
	Aug 1, 2013	Shijiazhuang	HEB	1050.0	7411.43	1.88%	11.3	6.1%	77.7%
8	Oct 9, 2013	Changchun	JL	752.7	10635.17	1.82%	11.3	2.0%	87.3%
	Oct 9, 2013	Nantong	JS	729.8	11047.17	1.49%	9.5	11.8%	88.3%
9	Mar 20, 2014	Anshan	LN	349.7	12000.70	0.38%	3.0	2.0%	87.1%
10	Apr 10, 2014	Huangshi	HUB	244.5	7473.41	0.19%	1.9	4.5%	72.4%
11	May 6, 2014	Weifang	SD	922.1	7670.88	0.91%	5.3	8.1%	66.3%
12	Jun 6, 2014	Harbin	HL	1012.2	7930.40	0.78%	4.5	1.8%	
13	Jul 10, 2014	Zhuzhou	HUN	393.5	7926.53	0.23%	1.8	0.6%	
14	Jul 23, 2014	Baoji	SAA	374.5	6605.40	0.14%	1.6	1.0%	
15	Aug 2, 2014	Yingkou	LN	232.5	10412.83	0.17%	1.7	4.1%	
16	Aug 6, 2014	Zhangzhou	FJ	493.0	7256.88	0.44%	2.8	2.8%	
17	Aug 21, 2014	Zhongshan	GD	317.4	13303.17	0.69%	3.9	3.6%	
18	Nov 6, 2014	Jilin	JL	429.1	9759.39	0.24%	1.9	1.1%	
19	Jan 1, 2015	Longyan	FJ	258.0	9177.67	0.41%	2.8	3.0%	
	Jan 4, 2015	Wuxi	JS	648.4	19913.77	0.73%	4.1	4.0%	
	Jan 4, 2015	Zibo	SD	458.6	13262.36	0.65%	3.8	12.4%	

Continued on next page

Table 3.4.1. Waves of Ucredit Office Expansion – Continued

Wave	Date	City	Prov.	Pop.	GDP p.c.	Freq.	# of lists per week	Funding prob. before	Funding prob. after
	Jan 5, 2015	Daqing	HL	282.6	23674.45	0.20%	1.7	2.1%	
	Jan 5, 2015	Jiangmen	GD	449.8	7115.53	0.47%	2.8	2.1%	
	Jan 5, 2015	Xuchang	HEN	429.7	6991.14	0.19%	1.7	1.4%	
	Jan 5, 2015	Xuzhou	JS	859.1	8261.34	0.60%	3.7	4.4%	
	Jan 6, 2015	Xianyang	SAA	494.2	6022.86	0.14%	1.7	17.3%	
20	Mar 2, 2015	Baoding	HEB	1022.9	4157.12	0.50%	3.2	4.2%	
	Mar 3, 2015	Fushun	LN	218.0	9838.14	0.13%	1.4	2.2%	
	Mar 3, 2015	Weihai	SD	280.6	14540.58	0.28%	2.2	5.0%	
	Mar 4, 2015	Dandong	LN	239.6	7394.30	0.19%	1.7	0.7%	
21	Apr 9, 2015	Taizhou	JS	463.4	10382.08	0.34%	2.4	0.8%	
	Apr 12, 2015	Yangzhou	JS	447.0	11640.30	0.27%	2.0	3.6%	
22	Apr 21, 2015	Songyuan	JL	282.9	9334.66	0.14%	1.5	4.8%	
23	May 1, 2015	Changde	HUN	607.2	5968.21	0.20%	1.9	2.8%	
	May 1, 2015	Changzhou	JS	469.2	14871.05	0.46%	2.9	6.9%	
	May 1, 2015	Jiaxing	ZJ	455.8	11049.28	0.51%	3.2	6.5%	
	May 1, 2015	Linyi	SD	1090.4	4896.27	0.82%	4.7	5.8%	
	May 1, 2015	Qujing	YN	597.4	4242.23	0.22%	1.7	1.3%	
	May 1, 2015	Shantou	GD	540.0	4639.72	0.52%	3.3	4.5%	
	May 1, 2015	Tangshan	HEB	747.4	13104.02	0.45%	2.9	3.4%	

§ Waves are defined by the week when Ucredit opens the first local office in a city. By Nov 27, 2012 when Renrendai initiated the UG program, Ucredit had already opened local offices in 7 cities–Beijing, Chengdu, Chongqing, Nanjing, Tianjin, Shenyang, and Suzhou, all included in wave 1.

† Data source of openning date: Website of Renrendai and Ucredit.

‡ Data source of population and GDP per capita: China City Statistical Yearbook 2014. The statistics record the levels in 2013. Population is in 10,000 people, and GDP per capita is in US dollars.

Table 3.4.2. Summary Statistics of Borrower Information

	Raw	In-use				
		All	Loan type		UG office opening	
			PP	UG	Pre	Post
Observations	254251	39779	19693	20086	13573	26206
Age	33.26	34.99	31.47	38.44	31.78	36.65
	(7.87)	(8.52)	(6.77)	(8.66)	(6.93)	(8.79)
Marriage Status						
NA	1.82%	0.01%	0.02%	0.00%	0.01%	0.01%
married	48.81%	59.06%	46.00%	71.88%	48.85%	64.35%
single	43.50%	34.71%	50.54%	19.20%	47.79%	27.94%
divorced	5.62%	5.90%	3.38%	8.37%	3.29%	7.25%
widowed	0.25%	0.32%	0.07%	0.56%	0.06%	0.45%
Education						
NA	9.57%	0.10%	0.20%	0.00%	0.19%	0.05%
high school	33.94%	35.55%	40.81%	30.40%	43.43%	31.47%
associate	37.56%	43.38%	37.45%	49.19%	37.15%	46.61%
bachelor	17.66%	19.58%	19.74%	19.42%	17.71%	20.54%
graduate	1.27%	1.39%	1.80%	0.99%	1.51%	1.33%
Monthly income						
NA	19.28%	0.26%	0.53%	0.00%	0.49%	0.15%
$< \$160$	0.27%	0.66%	0.54%	0.77%	0.50%	0.74%
$160-320	1.35%	1.66%	3.08%	0.26%	3.21%	0.85%
$320-800	27.87%	28.59%	38.79%	18.59%	40.74%	22.29%
$800-1600	27.85%	33.74%	30.45%	36.96%	29.08%	36.15%
$1600-3200	11.07%	15.77%	13.43%	18.07%	12.56%	17.43%
$3200-8000	7.17%	10.68%	7.68%	13.61%	7.65%	12.25%
$> \$8000$	5.15%	8.65%	5.50%	11.74%	5.77%	10.15%
Work experience (in years)						
NA	23.30%	0.38%	0.77%	0.00%	0.69%	0.22%
< 1	29.94%	26.88%	19.86%	33.75%	18.59%	31.17%
1-3	25.91%	39.83%	45.93%	33.85%	45.05%	37.13%
3-5	9.15%	15.00%	16.65%	13.39%	17.23%	13.85%
> 5	11.70%	17.91%	16.79%	19.01%	18.45%	17.63%
Company size (in employees)						
NA	22.75%	1.53%	3.08%	0.00%	2.98%	0.77%
< 10	28.93%	29.68%	21.57%	37.63%	22.42%	33.44%
10-100	27.86%	42.83%	40.02%	45.59%	40.20%	44.20%
100-500	8.74%	12.33%	15.49%	9.22%	15.13%	10.87%
> 500	11.72%	13.64%	19.83%	7.56%	19.27%	10.72%

Continued on next page

Table 3.4.2. Summary Statistics of Borrower Information – Continued

	Raw	In-use				
		All	Loan type		UG office opening	
			PP	UG	Pre	Post
Job type						
NA	0.04%	0.09%	0.19%	0.00%	0.21%	0.03%
others	1.51%	3.70%	7.47%	0.00%	7.75%	1.60%
working class	68.31%	61.37%	61.67%	61.08%	59.97%	62.10%
entrepreneur	26.18%	32.07%	25.10%	38.92%	26.75%	34.83%
e-dealer	3.95%	2.76%	5.58%	0.00%	5.32%	1.44%
House ownership						
no	63.71%	65.25%	60.12%	70.28%	57.83%	69.09%
yes	36.29%	34.75%	39.88%	29.72%	42.17%	30.91%
Mortgage						
no	83.14%	87.29%	87.14%	87.43%	86.97%	87.45%
yes	16.86%	12.71%	12.86%	12.57%	13.03%	12.55%
Car ownership						
no	80.39%	77.91%	75.41%	80.37%	74.10%	79.89%
yes	19.61%	22.09%	24.59%	19.63%	25.90%	20.11%
Auto loan						
no	95.23%	96.45%	95.24%	97.64%	95.14%	97.12%
yes	4.78%	3.55%	4.76%	2.37%	4.86%	2.88%

[§] Standard deviation of age is in the parantheses below.

Table 3.4.3. Impact of Credit Insurance on Total Number of Loan Listings

Week	(1)	(2)	(3)	(4)	(5)	(6)
-30	-0.349	0.335	-0.130	-0.0134	0.0209	0.0239
	(1.045)	(1.120)	(0.0978)	(0.106)	(0.0285)	(0.0356)
-29	0.966	0.799	0.134	0.0723	0.0200	0.0193
	(1.018)	(1.032)	(0.107)	(0.107)	(0.0278)	(0.0276)
-28	0.804	0.489	0.149*	0.146	0.0496	0.0416
	(1.089)	(1.109)	(0.0878)	(0.0920)	(0.0390)	(0.0334)
-27	0.337	0.832	-0.00795	0.0458	0.0665	0.0724
	(1.231)	(1.304)	(0.0947)	(0.106)	(0.0578)	(0.0488)
-26	0.286	0.418	0.175*	0.182*	0.0363	0.0385
	(1.172)	(1.203)	(0.102)	(0.103)	(0.0340)	(0.0319)
-25	0.569	0.254	-0.000712	-0.0658	0.0607	0.0534
	(1.134)	(1.140)	(0.0963)	(0.0976)	(0.0422)	(0.0377)
-24	1.721	1.983	0.260***	0.267***	0.0512	0.0539
	(1.174)	(1.227)	(0.0969)	(0.0947)	(0.0389)	(0.0374)
-23	0.592	0.592	0.121	0.159	0.0268	0.0304
	(1.078)	(1.161)	(0.0991)	(0.105)	(0.0338)	(0.0337)
-22	1.575	1.126	0.192*	0.193*	0.0587	0.0483
	(1.158)	(1.208)	(0.0987)	(0.104)	(0.0420)	(0.0374)
-21	0.475	0.491	0.0620	0.0379	0.0598	0.0591
	(1.172)	(1.168)	(0.113)	(0.112)	(0.0396)	(0.0370)
-20	1.230	1.439	0.234**	0.238**	0.0622*	0.0495
	(1.282)	(1.363)	(0.118)	(0.118)	(0.0370)	(0.0363)
-19	1.137	1.575	-0.000538	0.0260	0.0734	0.0778*
	(1.210)	(1.254)	(0.109)	(0.109)	(0.0486)	(0.0428)
-18	1.654	1.369	0.163*	0.160	0.0711*	0.0768*
	(1.177)	(1.191)	(0.0959)	(0.104)	(0.0431)	(0.0398)
-17	0.0568	1.168	0.0460	0.140	0.0739	0.0886**
	(1.235)	(1.308)	(0.109)	(0.117)	(0.0463)	(0.0398)
-16	0.405	0.731	0.00665	0.0317	0.0763	0.0639
	(1.183)	(1.211)	(0.105)	(0.109)	(0.0548)	(0.0465)
-15	2.003**	1.604	0.259***	0.206**	0.0558	0.0515
	(1.015)	(1.066)	(0.0973)	(0.100)	(0.0409)	(0.0372)
-14	1.280	1.056	0.106	0.115	0.0518	0.0539
	(1.138)	(1.194)	(0.105)	(0.109)	(0.0431)	(0.0386)
-13	1.605	1.452	0.00965	0.0209	0.0618	0.0515
	(1.243)	(1.304)	(0.150)	(0.138)	(0.0478)	(0.0457)
-12	1.270	1.105	0.0158	-0.0573	0.0569	0.0529
	(1.097)	(1.153)	(0.104)	(0.102)	(0.0473)	(0.0432)
-11	1.403	1.950	0.152*	0.240***	0.0400	0.0399

Continued on next page

Table 3.4.3. Impact of Credit Insurance on Total Number of Loan Listings – Continued

Week	(1)	(2)	(3)	(4)	(5)	(6)
	(1.095)	(1.199)	(0.0819)	(0.0853)	(0.0286)	(0.0282)
-10	2.175*	1.978	0.243**	0.251**	0.0727*	0.0717*
	(1.270)	(1.340)	(0.106)	(0.111)	(0.0373)	(0.0369)
-9	1.003	0.847	0.0963	0.0965	0.0545	0.0647*
	(1.186)	(1.231)	(0.107)	(0.112)	(0.0401)	(0.0390)
-8	-0.0418	0.661	-0.0157	0.0619	0.0248	0.0375
	(1.354)	(1.355)	(0.112)	(0.110)	(0.0286)	(0.0299)
-7	1.865	2.123	0.295**	0.293**	0.0421	0.0390
	(1.378)	(1.596)	(0.130)	(0.131)	(0.0349)	(0.0367)
-6	1.556	1.132	0.159	0.164	0.0829**	0.0686*
	(1.548)	(1.624)	(0.138)	(0.140)	(0.0394)	(0.0394)
-5	0.950	0.0176	0.122	0.0635	0.0487	0.0282
	(1.495)	(1.578)	(0.108)	(0.110)	(0.0346)	(0.0364)
-4	-0.330	0.221	-0.0624	-0.0661	0.0476	0.0622*
	(1.353)	(1.369)	(0.125)	(0.119)	(0.0366)	(0.0377)
-3	2.737**	2.836**	0.396***	0.382***	0.0766*	0.0896**
	(1.251)	(1.321)	(0.0951)	(0.0986)	(0.0423)	(0.0410)
-2	1.788	2.122	0.165	0.199*	0.0927*	0.104**
	(1.393)	(1.432)	(0.115)	(0.114)	(0.0514)	(0.0466)
-1	2.092*	3.000**	0.225**	0.320***	0.0700*	0.0916***
	(1.221)	(1.342)	(0.113)	(0.115)	(0.0357)	(0.0351)
0	3.948*	3.452*	0.410***	0.432***	0.272***	0.227***
	(2.076)	(2.055)	(0.129)	(0.130)	(0.0905)	(0.0687)
1	9.652***	8.628***	0.822***	0.708***	0.553***	0.541***
	(2.565)	(2.663)	(0.145)	(0.147)	(0.0654)	(0.0617)
2	16.82***	16.86***	1.183***	1.184***	0.712***	0.702***
	(3.762)	(3.689)	(0.172)	(0.162)	(0.0502)	(0.0509)
3	25.31***	25.16***	1.619***	1.632***	0.814***	0.804***
	(4.034)	(4.025)	(0.141)	(0.142)	(0.0334)	(0.0365)
4	29.69***	30.31***	1.636***	1.682***	0.804***	0.813***
	(3.570)	(3.594)	(0.149)	(0.154)	(0.0259)	(0.0260)
5	19.03***	19.54***	1.427***	1.451***	0.775***	0.781***
	(2.811)	(2.728)	(0.161)	(0.161)	(0.0266)	(0.0275)
6	21.08***	21.98***	1.325***	1.392***	0.788***	0.774***
	(3.406)	(3.361)	(0.169)	(0.167)	(0.0286)	(0.0334)
7	23.89***	23.67***	1.441***	1.388***	0.744***	0.723***
	(3.548)	(3.458)	(0.153)	(0.148)	(0.0317)	(0.0362)
8	28.54***	27.81***	1.655***	1.660***	0.769***	0.748***
	(3.701)	(3.569)	(0.164)	(0.154)	(0.0292)	(0.0312)
9	30.59***	31.43***	1.593***	1.694***	0.791***	0.789***
	(4.117)	(3.936)	(0.179)	(0.162)	(0.0270)	(0.0251)

Continued on next page

Table 3.4.3. Impact of Credit Insurance on Total Number of Loan Listings – Continued

Week	(1)	(2)	(3)	(4)	(5)	(6)
10	31.78***	31.49***	1.852***	1.846***	0.835***	0.826***
	(4.971)	(5.063)	(0.158)	(0.162)	(0.0180)	(0.0230)
11	26.28***	26.52***	1.488***	1.515***	0.785***	0.782***
	(4.083)	(4.079)	(0.175)	(0.165)	(0.0213)	(0.0242)
12	25.07***	24.95***	1.647***	1.633***	0.756***	0.744***
	(2.983)	(2.943)	(0.134)	(0.128)	(0.0223)	(0.0247)
13	24.74***	24.71***	1.611***	1.656***	0.758***	0.767***
	(2.991)	(3.053)	(0.139)	(0.143)	(0.0241)	(0.0255)
14	29.29***	29.55***	1.722***	1.684***	0.821***	0.825***
	(4.085)	(3.959)	(0.175)	(0.160)	(0.0238)	(0.0258)
15	33.52***	33.92***	1.739***	1.788***	0.832***	0.818***
	(4.328)	(4.280)	(0.180)	(0.183)	(0.0192)	(0.0217)
16	32.95***	33.04***	1.612***	1.703***	0.815***	0.795***
	(4.825)	(4.697)	(0.201)	(0.171)	(0.0264)	(0.0267)
17	29.10***	28.93***	1.681***	1.646***	0.768***	0.756***
	(3.526)	(3.511)	(0.148)	(0.149)	(0.0251)	(0.0256)
18	28.25***	29.26***	1.477***	1.528***	0.772***	0.776***
	(5.224)	(4.956)	(0.190)	(0.171)	(0.0297)	(0.0280)
19	31.68***	31.14***	1.699***	1.635***	0.788***	0.774***
	(4.812)	(4.871)	(0.165)	(0.167)	(0.0285)	(0.0327)
20	31.19***	30.79***	1.684***	1.682***	0.795***	0.780***
	(3.947)	(4.103)	(0.162)	(0.168)	(0.0243)	(0.0255)
21	27.12***	27.70***	1.474***	1.515***	0.744***	0.735***
	(4.146)	(4.037)	(0.182)	(0.179)	(0.0305)	(0.0307)
22	20.65***	21.27***	1.236***	1.352***	0.727***	0.694***
	(3.846)	(3.550)	(0.174)	(0.150)	(0.0370)	(0.0354)
23	32.63***	31.30***	1.670***	1.634***	0.744***	0.729***
	(4.717)	(4.578)	(0.177)	(0.172)	(0.0266)	(0.0271)
24	26.07***	24.72***	1.308***	1.189***	0.751***	0.733***
	(4.340)	(4.389)	(0.190)	(0.194)	(0.0325)	(0.0317)
25	30.51***	30.31***	1.380***	1.362***	0.718***	0.690***
	(5.141)	(5.153)	(0.205)	(0.211)	(0.0351)	(0.0363)
26	21.05***	23.48***	1.046***	1.219***	0.643***	0.678***
	(4.367)	(4.321)	(0.184)	(0.175)	(0.0462)	(0.0447)
27	28.26***	27.66***	1.375***	1.313***	0.728***	0.717***
	(5.387)	(5.211)	(0.176)	(0.168)	(0.0476)	(0.0447)
28	23.27***	24.44***	1.150***	1.251***	0.744***	0.731***
	(4.518)	(4.517)	(0.171)	(0.173)	(0.0310)	(0.0329)
29	33.29***	33.14***	1.533***	1.548***	0.752***	0.741***
	(5.147)	(5.163)	(0.167)	(0.177)	(0.0333)	(0.0342)
30	28.04***	27.57***	1.432***	1.376***	0.752***	0.748***

Continued on next page

Table 3.4.3. Impact of Credit Insurance on Total Number of Loan Listings – Continued

Week	(1)	(2)	(3)	(4)	(5)	(6)
	(4.544)	(4.516)	(0.158)	(0.162)	(0.0404)	(0.0403)
Observations	4,709	4,709	4,709	4,709	4,709	4,709
R^2	0.615	0.628	0.700	0.718	0.860	0.867
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes	Yes	Yes
Time-trend	Linear	Stochastic	Linear	Stochastic	Linear	Stochastic
Joint-test: pre	.69	.64	1.75	1.74	1.46	1.56
	[.898]	[.938]	[.007]	[.008]	[.051]	[.026]
Joint-test: post	37.75	39.36	34.41	36.72	268.72	246.53
	[0]	[0]	[0]	[0]	[0]	[0]

§ Numbers of loan listings are aggregates at city-week levels in column (1)-(2) and the logarithm of the city-week aggregates in column (3)-(4). Column (5)-(6) report weekly effects on shares of UG loans among total loan listings.

† Robust standard erros are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

‡ Joint-test: pre (post) tests jointly the 30 weekly effects before (since) Ucredit office opening. P-values are in brackets.

Table 3.4.4. Impact of Credit Insurance on Total Number of PP Loan Listings

Week	(1)	(2)	(3)	(4)
-30	-0.588	-0.307	-0.163*	-0.0733
	(0.368)	(0.416)	(0.0855)	(0.0936)
-29	0.267	0.155	0.0626	0.0109
	(0.461)	(0.452)	(0.0967)	(0.0989)
-28	0.271	0.249	0.0870	0.0977
	(0.329)	(0.326)	(0.0743)	(0.0753)
-27	-0.115	-0.00921	-0.0763	-0.0425
	(0.428)	(0.430)	(0.0820)	(0.0903)
-26	0.470	0.524	0.132	0.145
	(0.460)	(0.466)	(0.0863)	(0.0885)
-25	-0.524*	-0.699**	-0.0947	-0.152*
	(0.302)	(0.314)	(0.0830)	(0.0865)
-24	0.700	0.728	0.171*	0.168*
	(0.466)	(0.458)	(0.0883)	(0.0889)
-23	0.0690	0.205	0.0450	0.0912
	(0.540)	(0.512)	(0.0988)	(0.0992)
-22	0.468	0.461	0.106	0.115
	(0.398)	(0.423)	(0.0817)	(0.0878)
-21	-0.139	-0.119	0.00156	-0.0155
	(0.432)	(0.442)	(0.103)	(0.105)
-20	0.336	0.425	0.149	0.162*
	(0.481)	(0.478)	(0.0950)	(0.0923)
-19	0.0484	0.0906	-0.0840	-0.0754
	(0.471)	(0.465)	(0.101)	(0.0996)
-18	0.358	0.217	0.0725	0.0578
	(0.329)	(0.337)	(0.0821)	(0.0896)
-17	-0.480	-0.242	-0.0224	0.0511
	(0.366)	(0.380)	(0.0970)	(0.102)
-16	-0.508	-0.375	-0.0671	-0.0377
	(0.408)	(0.413)	(0.0928)	(0.0978)
-15	1.149**	0.990**	0.190**	0.148*
	(0.494)	(0.482)	(0.0796)	(0.0832)
-14	0.254	0.333	0.0296	0.0463
	(0.419)	(0.430)	(0.0996)	(0.103)
-13	0.459	0.490	-0.0824	-0.0632
	(0.697)	(0.668)	(0.140)	(0.128)
-12	-0.0773	-0.329	-0.0725	-0.142
	(0.400)	(0.408)	(0.0980)	(0.0963)
-11	0.0627	0.381	0.0583	0.133*
	(0.464)	(0.453)	(0.0801)	(0.0780)
-10	0.330	0.339	0.121	0.136

Continued on next page

Table 3.4.4. Impact of Credit Insurance on Total Number of PP Loan Listings – Continued

Week	(1)	(2)	(3)	(4)
	(0.366)	(0.382)	(0.0813)	(0.0896)
-9	0.267	0.196	0.0229	0.0240
	(0.410)	(0.422)	(0.0938)	(0.0983)
-8	-0.743	-0.429	-0.110	-0.0448
	(0.632)	(0.558)	(0.115)	(0.105)
-7	1.078*	1.166**	0.180	0.179
	(0.587)	(0.590)	(0.124)	(0.124)
-6	0.359	0.438	0.0511	0.0753
	(0.644)	(0.625)	(0.105)	(0.106)
-5	0.232	0.0142	0.0195	-0.0171
	(0.531)	(0.521)	(0.0980)	(0.0978)
-4	-0.0772	-0.255	-0.120	-0.143
	(0.579)	(0.566)	(0.117)	(0.114)
-3	1.280***	1.231***	0.264***	0.242***
	(0.426)	(0.438)	(0.0871)	(0.0935)
-2	0.691	0.827	0.0434	0.0679
	(0.561)	(0.548)	(0.0969)	(0.0953)
-1	0.858*	1.113**	0.107	0.179*
	(0.446)	(0.436)	(0.0975)	(0.0994)
0	1.376**	1.582***	0.202**	0.259***
	(0.617)	(0.601)	(0.0888)	(0.0911)
1	1.807***	1.489**	0.198**	0.109
	(0.581)	(0.592)	(0.0820)	(0.0871)
2	1.756***	1.759***	0.198*	0.197*
	(0.610)	(0.610)	(0.104)	(0.104)
3	1.023	1.116*	0.0502	0.0766
	(0.681)	(0.668)	(0.0966)	(0.0938)
4	1.435**	1.455**	0.132*	0.154**
	(0.599)	(0.600)	(0.0754)	(0.0782)
5	0.284	0.414	0.0225	0.0402
	(0.663)	(0.661)	(0.102)	(0.106)
6	-0.195	0.113	-0.0657	-0.00270
	(0.509)	(0.509)	(0.0867)	(0.0889)
7	2.111**	1.895**	0.0683	0.0236
	(0.855)	(0.800)	(0.0998)	(0.0987)
8	2.380***	2.413***	0.161	0.188*
	(0.714)	(0.686)	(0.107)	(0.106)
9	0.976	1.231*	0.0110	0.0883
	(0.621)	(0.629)	(0.0967)	(0.0985)
10	-0.397	-0.496	-0.0700	-0.0849
	(0.658)	(0.646)	(0.128)	(0.132)

Continued on next page

Table 3.4.4. Impact of Credit Insurance on Total Number of PP Loan Listings – Continued

Week	(1)	(2)	(3)	(4)
11	0.0505	0.0526	0.0834	0.0834
	(0.535)	(0.483)	(0.0907)	(0.0858)
12	1.115**	1.163**	0.248***	0.256***
	(0.512)	(0.496)	(0.0902)	(0.0911)
13	0.836	1.016*	0.137	0.205**
	(0.558)	(0.575)	(0.0932)	(0.0929)
14	0.0193	-0.258	-0.0207	-0.0962
	(0.505)	(0.528)	(0.105)	(0.115)
15	0.235	0.458	-0.0678	-0.0257
	(0.650)	(0.646)	(0.111)	(0.115)
16	0.942	1.250**	0.0992	0.196**
	(0.668)	(0.621)	(0.0999)	(0.0938)
17	1.874**	1.859**	0.218**	0.207**
	(0.749)	(0.726)	(0.0943)	(0.0944)
18	0.812	0.861	0.0763	0.0995
	(0.519)	(0.549)	(0.101)	(0.108)
19	0.943	0.786	0.158	0.105
	(0.871)	(0.883)	(0.111)	(0.115)
20	0.687	0.634	0.115	0.128
	(0.580)	(0.584)	(0.0955)	(0.102)
21	1.596**	1.735***	0.175	0.204*
	(0.635)	(0.631)	(0.108)	(0.113)
22	0.468	0.988	-0.0248	0.0866
	(0.663)	(0.663)	(0.107)	(0.107)
23	2.128***	2.064***	0.263**	0.263**
	(0.558)	(0.557)	(0.103)	(0.108)
24	0.354	-0.00571	-0.0300	-0.120
	(0.723)	(0.716)	(0.113)	(0.115)
25	1.994**	2.067**	0.0708	0.0809
	(0.893)	(0.879)	(0.130)	(0.133)
26	1.807***	2.085***	0.0601	0.145
	(0.659)	(0.620)	(0.115)	(0.114)
27	2.218***	2.012***	0.197**	0.133
	(0.702)	(0.697)	(0.0982)	(0.0975)
28	0.0306	0.484	-0.0991	0.0113
	(0.503)	(0.524)	(0.104)	(0.112)
29	2.701***	2.707***	0.235***	0.255***
	(0.640)	(0.627)	(0.0774)	(0.0802)
30	2.093***	1.921***	0.155*	0.124
	(0.568)	(0.541)	(0.0921)	(0.0929)
Observations	4,709	4,709	4,700	4,700

Continued on next page

Table 3.4.4. Impact of Credit Insurance on Total Number of PP Loan Listings – Continued

Week	(1)	(2)	(3)	(4)
R^2	0.670	0.685	0.582	0.604
City FE	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes
Time-trend	Linear	Stochastic	Linear	Stochastic
Joint-test: pre	1.4	1.37	1.45	1.37
	[.074]	[.087]	[.052]	[.088]
Joint-test: post	3.46	3.53	1.49	1.32
	[0]	[0]	[.042]	[.114]

§ Numbers of PP loan listings are aggregates at city-week levels in column (1)-(2) and the logarithm of the city-week aggregates in column (3)-(4).

† Robust standard erros are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

‡ Joint-test: pre (post) tests jointly the 30 weekly effects before (since) Ucredit office opening. P-values are in brackets.

Table 3.4.5. Impact of Credit Insurance on Average Funding Probability

Week	Per listed loan		Per listed PP loan		Per dollar of listed loans		Per dollar of listed PP loans	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
-30	0.0344	0.0392	0.0376	0.0382	0.0596	0.0601	0.0560	0.0516
	(0.0369)	(0.0392)	(0.0352)	(0.0368)	(0.0370)	(0.0386)	(0.0343)	(0.0359)
-29	0.0390	0.0298	0.0324	0.0187	0.0492	0.0429	0.0347	0.0218
	(0.0276)	(0.0250)	(0.0234)	(0.0205)	(0.0311)	(0.0298)	(0.0241)	(0.0220)
-28	0.0249	0.0343	0.0200	0.0305	0.0300	0.0371	0.0178	0.0212
	(0.0190)	(0.0230)	(0.0179)	(0.0239)	(0.0264)	(0.0292)	(0.0212)	(0.0245)
-27	0.0487	0.0406	0.0423	0.0309	0.0512*	0.0448	0.0355	0.0289
	(0.0321)	(0.0301)	(0.0307)	(0.0270)	(0.0294)	(0.0282)	(0.0271)	(0.0246)
-26	0.0206	0.0274	0.0182	0.0258	0.0387	0.0465	0.0258	0.0367
	(0.0194)	(0.0223)	(0.0195)	(0.0219)	(0.0249)	(0.0290)	(0.0212)	(0.0258)
-25	0.0224	0.0249	0.0211	0.0208	0.0208	0.0216	0.0130	0.0110
	(0.0195)	(0.0216)	(0.0187)	(0.0219)	(0.0221)	(0.0234)	(0.0200)	(0.0217)
-24	0.0382	0.0290	0.0367	0.0299	0.0363	0.0291	0.0250	0.0197
	(0.0282)	(0.0247)	(0.0280)	(0.0244)	(0.0278)	(0.0244)	(0.0259)	(0.0226)
-23	0.0382	0.0485*	0.0392*	0.0500*	0.0386	0.0467	0.0314	0.0342
	(0.0240)	(0.0279)	(0.0231)	(0.0275)	(0.0270)	(0.0323)	(0.0241)	(0.0287)
-22	0.000462	3.11e-05	0.000340	-0.00546	-0.00440	-0.00618	-0.00993	-0.0137
	(0.0205)	(0.0194)	(0.0204)	(0.0168)	(0.0162)	(0.0147)	(0.0133)	(0.0109)
-21	-0.00265	-0.00215	-0.00205	-0.00405	0.0133	0.0176	0.00820	0.00771
	(0.0164)	(0.0169)	(0.0166)	(0.0162)	(0.0253)	(0.0275)	(0.0233)	(0.0244)
-20	-0.0163*	-0.0185**	-0.0167*	-0.0196**	0.00275	-0.000216	-0.00368	-0.00688
	(0.00876)	(0.00736)	(0.00937)	(0.00781)	(0.0232)	(0.0208)	(0.0200)	(0.0177)
-19	0.00478	0.0120	0.00311	0.00452	0.0253	0.0318	0.0152	0.0160
	(0.0146)	(0.0170)	(0.0133)	(0.0144)	(0.0214)	(0.0232)	(0.0181)	(0.0192)
-18	0.0476	0.0551	0.0407	0.0453	0.0317	0.0368	0.0175	0.0202

Continued on next page

Table 3.4.5. Impact of Credit Insurance on Average Funding Probability – Continued

Week	Per loan		Per PP loan		Per $		Per $ of PP loan	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(0.0324)	(0.0338)	(0.0309)	(0.0322)	(0.0249)	(0.0244)	(0.0223)	(0.0223)
-17	0.0188	0.0234	0.0181	0.0180	0.0541	0.0482	0.0432	0.0282
	(0.0256)	(0.0258)	(0.0252)	(0.0231)	(0.0384)	(0.0317)	(0.0331)	(0.0233)
-16	0.0334	0.0239	0.0332	0.0207	0.0303	0.0220	0.0265	0.0207
	(0.0318)	(0.0262)	(0.0323)	(0.0263)	(0.0313)	(0.0258)	(0.0310)	(0.0269)
-15	0.0317	0.0336	0.0317	0.0362	0.0207	0.0264	0.0171	0.0192
	(0.0245)	(0.0254)	(0.0260)	(0.0301)	(0.0243)	(0.0270)	(0.0223)	(0.0251)
-14	0.0159	0.0191	0.0183	0.0139	0.00893	0.0121	0.00654	0.00335
	(0.0252)	(0.0248)	(0.0282)	(0.0254)	(0.0182)	(0.0192)	(0.0181)	(0.0173)
-13	0.0194	0.0228	0.0150	0.0267	0.00360	0.00281	-0.00190	0.00442
	(0.0319)	(0.0364)	(0.0270)	(0.0324)	(0.0163)	(0.0158)	(0.0140)	(0.0168)
-12	0.0261	0.0333	0.0184	0.0235	0.0296	0.0360	0.0171	0.0219
	(0.0305)	(0.0340)	(0.0273)	(0.0299)	(0.0317)	(0.0359)	(0.0281)	(0.0338)
-11	0.0162	0.00980	0.00894	-0.000903	-0.0138**	-0.0156**	-0.0182***	-0.0215***
	(0.0201)	(0.0192)	(0.0183)	(0.0162)	(0.00681)	(0.00628)	(0.00549)	(0.00429)
-10	-0.0213***	-0.0225***	-0.0233***	-0.0252***	0.000400	-0.00402	-0.00830	-0.0136
	(0.00767)	(0.00667)	(0.00657)	(0.00531)	(0.0219)	(0.0184)	(0.0166)	(0.0127)
-9	0.0121	0.0195	0.00974	0.0192	0.0271	0.0366	0.0179	0.0317
	(0.0201)	(0.0228)	(0.0201)	(0.0221)	(0.0272)	(0.0310)	(0.0258)	(0.0323)
-8	-0.0133	-0.00456	-0.0151	-0.00812	-0.00517	0.00362	-0.00998	0.000104
	(0.0109)	(0.0161)	(0.00989)	(0.0128)	(0.0130)	(0.0180)	(0.0110)	(0.0153)
-7	-0.00651	-0.00897	-0.0101	-0.0104	0.0576	0.0554	0.0424	0.0399
	(0.0117)	(0.0106)	(0.0114)	(0.0121)	(0.0410)	(0.0409)	(0.0357)	(0.0380)
-6	0.0143	0.0241	0.00998	0.0145	0.0216	0.0344	0.0115	0.0206
	(0.0206)	(0.0249)	(0.0173)	(0.0205)	(0.0270)	(0.0331)	(0.0226)	(0.0281)

Continued on next page

Table 3.4.5. Impact of Credit Insurance on Average Funding Probability – Continued

Week	Per loan		Per PP loan		Per $		Per $ of PP loan	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
-5	0.0781	0.0552	0.0686	0.0440	0.0999*	0.0721	0.0749	0.0472
	(0.0546)	(0.0457)	(0.0527)	(0.0454)	(0.0578)	(0.0497)	(0.0552)	(0.0476)
-4	-0.0168*	-0.0155	-0.0173*	-0.0184**	-0.0125	-0.0116	-0.0159**	-0.0169**
	(0.00911)	(0.0101)	(0.00933)	(0.00894)	(0.00913)	(0.00984)	(0.00794)	(0.00763)
-3	0.0154	0.0223	0.00799	0.0100	0.0370	0.0408	0.0240	0.0263
	(0.0274)	(0.0303)	(0.0219)	(0.0232)	(0.0467)	(0.0471)	(0.0368)	(0.0373)
-2	0.0252	0.0384	0.0176	0.0371	0.0555	0.0629	0.0398	0.0564
	(0.0296)	(0.0377)	(0.0257)	(0.0337)	(0.0555)	(0.0622)	(0.0464)	(0.0569)
-1	0.0112	0.00947	0.00415	-0.00217	-0.000843	0.000706	-0.00724	-0.0112
	(0.0221)	(0.0222)	(0.0200)	(0.0172)	(0.0186)	(0.0204)	(0.0151)	(0.0131)
0	0.126**	0.110**	-0.00278	-0.00303	0.148**	0.129**	0.0215	0.0143
	(0.0589)	(0.0522)	(0.0160)	(0.0173)	(0.0591)	(0.0528)	(0.0281)	(0.0279)
1	0.398***	0.368***	0.0330	0.0227	0.432***	0.419***	0.0621	0.0577
	(0.0659)	(0.0633)	(0.0280)	(0.0255)	(0.0621)	(0.0584)	(0.0623)	(0.0622)
2	0.544***	0.545***	0.0184	0.0207	0.504***	0.513***	0.00512	0.00734
	(0.0601)	(0.0624)	(0.0260)	(0.0285)	(0.0545)	(0.0580)	(0.0172)	(0.0208)
3	0.699***	0.674***	0.0372	0.0259	0.698***	0.655***	0.0597	0.0373
	(0.0484)	(0.0547)	(0.0296)	(0.0265)	(0.0490)	(0.0569)	(0.0498)	(0.0392)
4	0.675***	0.691***	0.00806	0.00858	0.658***	0.673***	0.0263	0.0268
	(0.0363)	(0.0324)	(0.0206)	(0.0224)	(0.0356)	(0.0316)	(0.0307)	(0.0319)
5	0.722***	0.712***	0.0881***	0.0689**	0.670***	0.665***	0.0587**	0.0397
	(0.0279)	(0.0312)	(0.0328)	(0.0294)	(0.0325)	(0.0375)	(0.0289)	(0.0293)
6	0.637***	0.632***	0.0111	0.00865	0.617***	0.608***	0.0262	0.0266
	(0.0398)	(0.0404)	(0.0385)	(0.0370)	(0.0376)	(0.0408)	(0.0457)	(0.0491)
7	0.646***	0.635***	0.00361	0.00628	0.614***	0.621***	-0.0171**	-0.0151

Continued on next page

Table 3.4.5. Impact of Credit Insurance on Average Funding Probability – Continued

Week	Per loan		Per PP loan		Per $		Per $ of PP loan	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(0.0416)	(0.0414)	(0.0221)	(0.0236)	(0.0379)	(0.0400)	(0.00803)	(0.0106)
8	0.692***	0.680***	0.0501	0.0464	0.667***	0.632***	0.0621	0.0500
	(0.0369)	(0.0389)	(0.0582)	(0.0562)	(0.0357)	(0.0413)	(0.0655)	(0.0557)
9	0.682***	0.746***	-0.0164	-0.0140	0.657***	0.702***	-0.0128	-0.0110
	(0.0433)	(0.0257)	(0.0109)	(0.0130)	(0.0436)	(0.0375)	(0.0121)	(0.0142)
10	0.782***	0.772***	-0.00885	-0.00691	0.789***	0.792***	-0.0237***	-0.0223***
	(0.0215)	(0.0270)	(0.0206)	(0.0234)	(0.0226)	(0.0259)	(0.00321)	(0.00478)
11	0.708***	0.715***	0.0385	0.0336	0.720***	0.713***	0.0746	0.0658
	(0.0288)	(0.0349)	(0.0320)	(0.0308)	(0.0298)	(0.0380)	(0.0636)	(0.0670)
12	0.723***	0.669***	0.0349	0.0164	0.693***	0.626***	-0.00321	-0.0161*
	(0.0232)	(0.0329)	(0.0371)	(0.0293)	(0.0274)	(0.0394)	(0.0157)	(0.00874)
13	0.728***	0.733***	0.0297	0.0584	0.728***	0.718***	0.0194	0.0388
	(0.0262)	(0.0256)	(0.0330)	(0.0464)	(0.0276)	(0.0282)	(0.0257)	(0.0358)
14	0.714***	0.750***	-0.000153	-0.00571	0.704***	0.748***	0.00816	0.00957
	(0.0400)	(0.0281)	(0.0246)	(0.0206)	(0.0382)	(0.0293)	(0.0284)	(0.0285)
15	0.753***	0.766***	0.0778*	0.112**	0.754***	0.770***	0.0537	0.0848*
	(0.0241)	(0.0258)	(0.0400)	(0.0473)	(0.0255)	(0.0258)	(0.0333)	(0.0447)
16	0.687***	0.686***	0.0255	0.00843	0.694***	0.693***	0.00848	-0.00351
	(0.0461)	(0.0437)	(0.0262)	(0.0205)	(0.0443)	(0.0446)	(0.0209)	(0.0165)
17	0.694***	0.690***	0.00319	0.0129	0.684***	0.651***	0.000330	0.00233
	(0.0356)	(0.0330)	(0.0209)	(0.0245)	(0.0345)	(0.0340)	(0.0171)	(0.0176)
18	0.667***	0.700***	0.0150	0.0177	0.654***	0.675***	-0.00951	-0.00862
	(0.0403)	(0.0342)	(0.0274)	(0.0309)	(0.0397)	(0.0362)	(0.0112)	(0.0131)
19	0.712***	0.686***	0.0680*	0.0461	0.718***	0.721***	0.0366	0.0406
	(0.0339)	(0.0380)	(0.0360)	(0.0332)	(0.0339)	(0.0340)	(0.0322)	(0.0368)

Continued on next page

Table 3.4.5. Impact of Credit Insurance on Average Funding Probability – Continued

Week	Per loan		Per PP loan		Per $		Per $ of PP loan	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
20	0.709***	0.700***	-0.0298***	-0.0300***	0.676***	0.662***	-0.0258***	-0.0262***
	(0.0308)	(0.0302)	(0.00207)	(0.00208)	(0.0321)	(0.0345)	(0.00211)	(0.00217)
21	0.637***	0.626***	0.0189	0.0112	0.629***	0.594***	-0.0136*	-0.0186***
	(0.0423)	(0.0428)	(0.0243)	(0.0239)	(0.0425)	(0.0461)	(0.00743)	(0.00546)
22	0.635***	0.662***	0.0149	0.0506	0.658***	0.655***	0.00209	0.0487
	(0.0473)	(0.0393)	(0.0266)	(0.0471)	(0.0445)	(0.0370)	(0.0182)	(0.0413)
23	0.679***	0.664***	0.0613	0.0640	0.685***	0.664***	0.107	0.121
	(0.0383)	(0.0390)	(0.0540)	(0.0558)	(0.0343)	(0.0369)	(0.0880)	(0.0979)
24	0.653***	0.597***	0.0723	0.0512	0.662***	0.646***	0.113*	0.119**
	(0.0451)	(0.0525)	(0.0546)	(0.0454)	(0.0388)	(0.0421)	(0.0658)	(0.0602)
25	0.630***	0.552***	0.0241	0.00771	0.637***	0.547***	-0.00196	-0.0145
	(0.0439)	(0.0531)	(0.0264)	(0.0214)	(0.0406)	(0.0515)	(0.0152)	(0.00973)
26	0.490***	0.602***	0.00459	0.0225	0.515***	0.596***	-0.0123	-0.00710
	(0.0598)	(0.0567)	(0.0208)	(0.0326)	(0.0540)	(0.0524)	(0.0111)	(0.0168)
27	0.610***	0.619***	0.0261	0.0235	0.639***	0.639***	0.0375	0.0288
	(0.0475)	(0.0458)	(0.0267)	(0.0253)	(0.0447)	(0.0439)	(0.0360)	(0.0358)
28	0.674***	0.677***	0.0567	0.0585	0.688***	0.693***	0.0694	0.0841
	(0.0457)	(0.0461)	(0.0479)	(0.0538)	(0.0442)	(0.0431)	(0.0662)	(0.0757)
29	0.664***	0.627***	0.000375	-0.00439	0.660***	0.641***	-0.0177**	-0.0167*
	(0.0442)	(0.0497)	(0.0235)	(0.0206)	(0.0427)	(0.0467)	(0.00785)	(0.00918)
30	0.658***	0.641***	-0.0233***	-0.0223**	0.670***	0.625***	-0.0243***	-0.0248***
	(0.0459)	(0.0503)	(0.00696)	(0.00889)	(0.0454)	(0.0510)	(0.00279)	(0.00289)
Obs	4,709	4,709	4,700	4,700	4,709	4,709	4,700	4,700
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Continued on next page

Table 3.4.5. Impact of Credit Insurance on Average Funding Probability – Continued

Week	Per loan		Per PP loan		Per $		Per $ of PP loan	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Time-trend	Linear	Stoch	Linear	Stoch	Linear	Stoch	Linear	Stoch
Joint-test:								
pre	37.12	37.68	33.39	34.51	31.13	33.32	28.03	32.76
	[.174]	[.158]	[.306]	[.261]	[.409]	[.309]	[.569]	[.333]
post	691.48	721.37	2813.69	1488.34	57.07	49.09	49.72	45.99
	[0]	[0]	[0]	[0]	[.001]	[.011]	[.01]	[.024]

§ Funding probabilities are measured per loan at city-week averages in column (1)-(4), and per dollar at city-week averages in column (5)-(8).

† Robust standard erros are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

‡ Joint-test: pre (post) tests jointly the 30 weekly effects before (since) Ucredit office opening. P-values are in brackets.

Table 3.4.6. Impacts of Credit Insurance on Average Funding Speed and Bid Amount

Week	Per funded loan		Per funded PP loan		Per funded loan		Per funded PP loan	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
-30	-166.7***	-109.0***	-151.0**	-107.9*	37.37	7.130	42.75	-13.63
	(47.92)	(40.32)	(70.00)	(63.86)	(52.44)	(21.34)	(48.36)	(32.14)
-29	-59.48	-98.33**	-42.07	-79.59	-18.43	-12.61	-3.522	-11.68
	(40.05)	(44.70)	(41.38)	(52.03)	(21.63)	(24.22)	(23.77)	(28.29)
-28	-94.25***	-46.37	-58.11	3.534	43.00**	41.96	39.81	33.92
	(26.79)	(32.12)	(36.14)	(44.66)	(20.62)	(25.83)	(27.06)	(38.87)
-27	-41.84	-30.09	-19.52	1.113	15.67	26.19	9.433	12.15
	(55.14)	(58.75)	(65.99)	(76.66)	(17.68)	(19.76)	(27.97)	(31.21)
-26	-118.3***	-114.6***	-63.05*	-60.15	11.42	-1.749	1.147	-11.68
	(25.70)	(32.87)	(34.86)	(41.36)	(21.05)	(21.84)	(23.64)	(30.77)
-25	28.39	12.32	68.54	70.03	4.214	22.87	24.62	-2.907
	(59.32)	(61.45)	(61.66)	(68.88)	(33.87)	(38.30)	(39.93)	(38.31)
-24	-65.97	-62.25	-19.40	26.18	-5.213	27.31	6.053	14.08
	(41.35)	(50.49)	(47.36)	(56.91)	(22.52)	(25.42)	(32.13)	(35.88)
-23	-114.3***	-111.7***	-71.12*	-37.05	8.757	3.863	9.347	-2.899
	(36.23)	(42.46)	(39.03)	(54.44)	(28.98)	(27.54)	(34.20)	(33.09)
-22	38.94	68.40	51.59	76.89	103.8	94.59	77.56	67.91
	(82.20)	(75.64)	(106.5)	(114.1)	(64.76)	(88.51)	(60.17)	(92.39)
-21	-74.40	-75.23	-22.69	-20.24	-96.00***	-113.6***	-143.6**	-168.4***
	(60.85)	(64.90)	(60.89)	(82.39)	(33.40)	(30.21)	(61.13)	(63.98)
-20	-56.03*	-50.32*	-47.98	-16.23	105.6***	164.7***	132.3***	177.3***
	(28.74)	(28.75)	(47.44)	(43.45)	(28.38)	(32.34)	(39.96)	(46.85)
-19	-29.42	-7.398	0.192	39.19	-21.85	40.39	-8.606	37.24
	(75.81)	(80.64)	(73.49)	(81.62)	(42.73)	(45.07)	(52.31)	(52.01)
-18	-46.23	-52.58	-47.50	-40.90	-16.86	-3.338	12.96	6.721
	(58.78)	(73.24)	(45.05)	(62.18)	(43.60)	(48.48)	(46.29)	(47.53)
-17	-97.08*	-21.22	-13.45	57.88	296.6**	222.0**	317.3**	211.1
	(49.61)	(53.53)	(49.58)	(57.52)	(122.4)	(100.9)	(139.4)	(135.6)
-16	-7.591	-55.15	2.964	-30.82	37.37	79.38	147.9*	174.0
	(51.51)	(88.72)	(49.04)	(94.59)	(51.59)	(162.5)	(87.24)	(144.7)
-15	-65.31	-44.86	-58.33	-42.27	37.04	36.46	55.85	32.98
	(50.34)	(48.64)	(52.83)	(57.59)	(35.99)	(73.00)	(43.19)	(71.23)
-14	20.79	21.56	31.72	56.06	-62.06*	-83.81*	-51.95	-74.22
	(109.1)	(124.4)	(124.5)	(140.0)	(34.02)	(43.30)	(40.27)	(63.30)
-13	-10.75	-15.03	4.831	-8.868	-93.17**	-93.67*	-103.6*	-134.6
	(111.6)	(122.0)	(93.53)	(93.02)	(45.38)	(50.31)	(61.56)	(95.33)
-12	-105.3***	-113.8***	-68.94	-32.84	127.6	187.2	155.1	153.1
	(36.19)	(39.78)	(56.49)	(71.17)	(152.8)	(182.0)	(143.5)	(180.3)
-11	-81.73*	-83.09*	-39.00	-37.66	-8.799	43.26	-9.669	0.860
	(45.47)	(48.15)	(48.92)	(53.14)	(29.91)	(40.98)	(37.32)	(44.86)

Continues on next page

Table 3.4.6. Impacts of Credit Insurance on Average Funding Speed and Bid Amount – Continued

Week	Per funded loan		Per funded PP loan		Per funded loan		Per funded PP loan	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
-10	-130.4***	-126.1***	-51.43	6.313	49.10**	36.34	23.56	-172.0
	(22.10)	(20.92)	(43.51)	(56.80)	(23.61)	(26.39)	(47.10)	(115.7)
-9	-8.306	6.884	-0.472	-13.07	-12.18	-65.76	5.515	-39.23
	(82.33)	(80.99)	(66.49)	(66.83)	(35.51)	(51.05)	(51.45)	(64.62)
-8	-32.38	-29.86	-49.23	-5.166	-108.1***	-93.45***	-35.16	-94.00*
	(79.60)	(81.43)	(50.92)	(49.32)	(20.34)	(22.09)	(31.05)	(48.63)
-7	-130.7***	-133.9***	-82.88	-57.06	158.7*	125.8**	164.7***	119.4
	(25.06)	(22.23)	(55.63)	(65.50)	(81.83)	(62.02)	(60.39)	(72.91)
-6	-17.68	6.217	-26.30	-16.02	152.9	86.05	226.2**	150.1*
	(54.09)	(45.54)	(50.91)	(47.92)	(101.6)	(94.49)	(90.30)	(84.10)
-5	-70.79	-67.56	-32.65	3.082	28.45	11.55	95.43*	85.60
	(64.00)	(67.79)	(70.59)	(62.83)	(53.46)	(39.58)	(54.08)	(58.93)
-4	-72.96	-54.35	-41.99	-37.46	-84.80***	-19.28	-69.57*	-31.80
	(47.43)	(64.77)	(58.75)	(89.48)	(30.63)	(34.77)	(41.88)	(67.39)
-3	129.8	133.6	95.02	68.03	-100.5**	-81.07*	-115.3**	-109.2
	(127.4)	(152.5)	(89.80)	(106.2)	(40.45)	(46.42)	(54.69)	(79.82)
-2	118.5*	137.5**	118.0**	164.3**	-73.34***	-43.93*	-86.05***	-102.5***
	(64.02)	(54.77)	(58.58)	(69.88)	(25.42)	(25.80)	(29.66)	(35.77)
-1	-0.168	22.84	-0.957	36.86	-28.85	-48.52	-56.22	-51.31
	(96.91)	(101.7)	(72.52)	(81.93)	(60.15)	(93.02)	(72.02)	(120.1)
0	-98.35***	-121.2***	-51.36	-60.98	-49.85	-38.35	23.13	-6.494
	(32.14)	(30.22)	(35.00)	(41.77)	(38.64)	(81.36)	(48.66)	(118.1)
1	-146.6***	-150.5***	7.331	19.45	3.058	28.53	-45.90	-77.15
	(24.39)	(22.97)	(88.78)	(83.65)	(28.72)	(33.35)	(46.32)	(61.68)
2	-152.8***	-153.9***	-22.79	-64.35	45.88**	75.77***	11.50	46.61
	(24.47)	(23.49)	(48.55)	(65.47)	(20.12)	(23.98)	(30.30)	(34.37)
3	-154.4***	-153.2***	-115.9	-94.66	32.40*	11.89	56.42	-67.97
	(24.69)	(23.36)	(71.28)	(77.01)	(19.06)	(19.48)	(51.74)	(100.5)
4	-154.2***	-152.8***	28.27	15.24	32.06	34.95	-19.71	42.11
	(24.91)	(23.47)	(45.74)	(66.04)	(21.32)	(22.39)	(37.92)	(41.65)
5	-156.3***	-155.3***	-94.89**	-81.13*	38.23*	30.62	-8.432	-58.08
	(25.16)	(23.74)	(45.97)	(48.29)	(21.12)	(23.08)	(41.52)	(47.80)
6	-157.9***	-158.1***	-15.02	-54.37	57.38***	53.39**	-3,864***	-3,831***
	(25.39)	(23.99)	(77.90)	(108.2)	(21.79)	(23.20)	(37.40)	(39.21)
7	-158.7***	-158.8***	-9.377	4.381	47.63**	53.90**	-37.16	-13.51
	(25.64)	(24.26)	(73.77)	(77.11)	(20.38)	(22.08)	(32.52)	(38.42)
8	-159.6***	-158.9***	-35.02	-46.40	60.87***	48.37**	-33.21	-54.07
	(25.88)	(24.50)	(47.04)	(53.87)	(20.53)	(20.22)	(31.22)	(37.52)
9	-162.0***	-160.8***	-62.72	-77.72	38.84*	19.62	96.47	42.58

Continues on next page

Table 3.4.6. Impacts of Credit Insurance on Average Funding Speed and Bid Amount – Continued

Week	Per funded loan		Per funded PP loan		Per funded loan		Per funded PP loan	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(26.11)	(24.67)	(44.78)	(49.40)	(20.75)	(21.76)	(109.7)	(134.6)
10	-164.1***	-163.8***	178.5***	160.4*	55.75***	64.15***	-357.6***	-40.78
	(26.39)	(24.91)	(65.44)	(97.10)	(21.52)	(20.43)	(98.35)	(39.52)
11	-164.4***	-165.1***	92.18	99.93	55.75***	50.37**	-21.09	2.680
	(26.65)	(25.21)	(67.21)	(73.39)	(19.63)	(19.99)	(31.91)	(37.33)
12	-167.4***	-168.0***	-100.7	-178.1	55.90**	57.18**	7.487	5.396
	(26.88)	(25.46)	(94.00)	(114.8)	(24.71)	(23.82)	(65.03)	(90.12)
13	-168.7***	-169.7***	-125.8*	-117.9	45.85**	27.70	149.7	28.86
	(27.15)	(25.73)	(66.83)	(89.70)	(21.38)	(23.54)	(100.8)	(64.07)
14	-171.1***	-172.2***	57.43	143.8**	75.19***	60.66***	-39.43	-101.3*
	(27.41)	(25.98)	(51.92)	(65.14)	(20.89)	(22.36)	(46.15)	(61.28)
15	-171.3***	-172.0***	87.24	163.4*	51.69**	56.75**	-12.79	-66.79
	(27.68)	(26.21)	(87.95)	(96.34)	(23.87)	(22.16)	(64.96)	(100.3)
16	-174.1***	-175.2***	-99.55	-146.3	46.47**	61.25***	-171.4**	-146.7*
	(27.95)	(26.53)	(98.51)	(97.41)	(21.56)	(21.23)	(73.25)	(84.03)
17	-174.5***	-176.0***	29.86	-49.84	68.68***	61.12***	98.09	94.76
	(28.20)	(26.79)	(55.41)	(57.38)	(24.63)	(23.08)	(116.1)	(105.8)
18	-177.2***	-178.2***	-36.06	-20.06	56.68**	42.44**	44.46	-2.867
	(28.49)	(26.97)	(63.79)	(68.44)	(23.72)	(21.51)	(82.14)	(98.55)
19	-177.6***	-179.1***	-11.62	15.31	100.1***	95.30***	-79.16**	-88.55
	(28.77)	(27.31)	(57.96)	(60.00)	(21.16)	(21.12)	(32.94)	(62.50)
20	-180.7***	-182.5***			105.0***	92.30***		
	(29.02)	(27.54)			(22.61)	(22.23)		
21	-181.9***	-184.1***	-42.00	-147.4	43.37*	76.20***	-43.21	-24.74
	(29.30)	(28.00)	(66.96)	(108.8)	(23.25)	(22.29)	(53.30)	(54.65)
22	-183.7***	-184.7***	-169.2***	-168.9***	75.97***	48.02*	-56.30	-22.83
	(29.57)	(28.16)	(47.55)	(60.38)	(27.00)	(27.67)	(48.44)	(42.58)
23	-184.2***	-185.3***	-4.392	174.7	45.70**	52.20**	-18.76	-66.58
	(29.86)	(28.39)	(103.8)	(137.2)	(23.08)	(23.52)	(33.81)	(47.99)
24	-185.4***	-188.0***	44.79	104.2	82.25***	78.96***	5.209	-28.92
	(30.15)	(28.66)	(85.16)	(67.58)	(24.40)	(23.42)	(69.29)	(84.07)
25	-185.8***	-188.6***	21.49	-14.09	66.24***	89.07***	15.73	-82.28
	(30.43)	(28.99)	(84.42)	(84.19)	(24.60)	(22.74)	(48.09)	(63.43)
26	-187.7***	-191.1***	40.22	135.9*	86.54***	57.20**	-93.57***	-119.7**
	(30.71)	(29.36)	(61.01)	(72.83)	(25.69)	(28.40)	(34.99)	(51.22)
27	-191.2***	-194.5***	-138.5*	-96.02	71.54***	57.75**	-16.21	-25.51
	(30.99)	(29.55)	(74.13)	(86.18)	(25.84)	(25.77)	(45.03)	(53.09)
28	-191.2***	-193.1***	25.68	55.52	106.9***	91.68***	55.39	108.9
	(31.29)	(29.78)	(85.40)	(124.6)	(26.88)	(26.45)	(75.84)	(128.0)

Continues on next page

Table 3.4.6. Impacts of Credit Insurance on Average Funding Speed and Bid Amount – Continued

Week	Per funded loan		Per funded PP loan		Per funded loan		Per funded PP loan	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
29	-195.0***	-198.7***	-55.12	-54.56	77.89***	77.66***	-63.68	-174.9*
	(31.55)	(30.11)	(69.76)	(102.2)	(23.57)	(22.57)	(50.43)	(105.7)
30	-195.6***	-200.3***	-101.4	-159.4**	63.61**	64.72**	-44.14	24.15
	(31.79)	(30.51)	(63.17)	(73.00)	(25.64)	(25.84)	(34.29)	(65.65)
Obs	1,024	1,024	537	537	1,024	1,024	537	537
R^2	0.483	0.524	0.414	0.551	0.456	0.637	0.569	0.671
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time-trend	Linear	Stoch	Linear	Stoch	Linear	Stoch	Linear	Stoch
Joint-test:								
pre	3.45	3.72	.89	.87	10.23	4.12	8.68	3.23
	[0]	[0]	[.631]	[.672]	[0]	[0]	[0]	[0]
post	2.17	2.15	3	2.78	2.16	2.13	717.02	579.45
	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]

§ Bidding speed is defined as the time difference (in hours) per loan between the first and last bid at city-week unweighted averages in column (1) -(4). Bid amounts are measured in US dollars per bid at city-week unweighted averages in column (5) -(8).

† Robust standard erros are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

‡ Joint-test: pre (post) tests jointly the 30 weekly effects before (since) Ucredit office opening. P-values are in brackets.

Table 3.4.7. Impact of Credit Insurance on Average Funding Probability (with Borroower Information)

Week	Per listed loan				Per listed PP loan			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
-30	0.00489	0.00543	-0.00464	0.00154	0.00913	0.0103	0.00701	0.00833
	(0.0106)	(0.0117)	(0.0222)	(0.0249)	(0.0151)	(0.0177)	(0.0136)	(0.0164)
-29	0.0178	0.0272*	0.0440	0.0538*	0.0295*	0.0298	0.0242	0.0226
	(0.0120)	(0.0152)	(0.0281)	(0.0320)	(0.0169)	(0.0190)	(0.0149)	(0.0166)
-28	0.0260*	0.0262*	0.0590**	0.0501*	0.0221	0.0179	0.0102	0.00141
	(0.0140)	(0.0146)	(0.0291)	(0.0303)	(0.0148)	(0.0159)	(0.0107)	(0.0108)
-27	0.0230	0.0181	0.0344	0.0255	0.00852	0.0105	0.000723	0.00261
	(0.0153)	(0.0152)	(0.0289)	(0.0291)	(0.0135)	(0.0157)	(0.0103)	(0.0117)
-26	0.0115	0.0161	0.0333	0.0531	0.0112	0.0177	0.0129	0.0206
	(0.0119)	(0.0138)	(0.0294)	(0.0353)	(0.0140)	(0.0181)	(0.0125)	(0.0159)
-25	0.0227	0.0416**	0.0433	0.0769*	0.0363*	0.0431*	0.0310*	0.0419*
	(0.0142)	(0.0196)	(0.0323)	(0.0408)	(0.0195)	(0.0241)	(0.0178)	(0.0220)
-24	0.0302**	0.0270*	0.0469*	0.0485*	0.0359**	0.0314*	0.0227*	0.0241*
	(0.0141)	(0.0138)	(0.0272)	(0.0282)	(0.0174)	(0.0182)	(0.0127)	(0.0146)
-23	0.0245*	0.0253	0.0305	0.0306	0.0261	0.0329	0.0112	0.0114
	(0.0140)	(0.0156)	(0.0285)	(0.0324)	(0.0169)	(0.0212)	(0.0124)	(0.0147)
-22	0.00310	-0.00380	0.000732	-0.0139	-0.00567	-0.00950	-0.00682	-0.0106
	(0.0118)	(0.00867)	(0.0290)	(0.0235)	(0.0116)	(0.0110)	(0.0118)	(0.0120)
-21	-0.00432	0.000936	-0.0164	-0.00751	-0.00435	-0.00493	-0.00615	-0.00979
	(0.00965)	(0.0127)	(0.0247)	(0.0290)	(0.0141)	(0.0156)	(0.0128)	(0.0127)
-20	-0.00388	-0.00433	-0.0163	-0.0165	-0.00498	-0.0104	-0.0133	-0.0139
	(0.00962)	(0.00933)	(0.0216)	(0.0219)	(0.0136)	(0.0120)	(0.0105)	(0.0115)
-19	0.0280	0.0237	0.0311	0.0301	0.0197	0.0156	0.00478	0.00114
	(0.0173)	(0.0156)	(0.0308)	(0.0292)	(0.0172)	(0.0171)	(0.0126)	(0.0128)
-18	0.0374**	0.0370**	0.0775**	0.0827**	0.0254	0.0434*	0.0237	0.0314
	(0.0182)	(0.0188)	(0.0348)	(0.0367)	(0.0182)	(0.0258)	(0.0158)	(0.0198)
-17	0.0190	0.0238	0.0387	0.0485	0.0233	0.0140	0.0153	0.00123
	(0.0161)	(0.0167)	(0.0301)	(0.0315)	(0.0195)	(0.0167)	(0.0134)	(0.0111)
-16	0.00406	0.0153	0.0195	0.0423	0.0120	0.0191	0.0105	0.0126
	(0.0127)	(0.0174)	(0.0323)	(0.0387)	(0.0197)	(0.0240)	(0.0164)	(0.0192)
-15	0.0282*	0.0377**	0.0705**	0.0816**	0.0386**	0.0480**	0.0305**	0.0348*
	(0.0151)	(0.0167)	(0.0329)	(0.0339)	(0.0197)	(0.0233)	(0.0155)	(0.0183)
-14	0.00572	0.00701	0.0271	0.0219	0.00551	0.00219	0.00527	-0.000682
	(0.0131)	(0.0138)	(0.0326)	(0.0322)	(0.0164)	(0.0170)	(0.0152)	(0.0147)
-13	0.00339	-0.00410	0.00346	-0.0113	-0.00140	0.00406	0.000842	0.0107
	(0.0136)	(0.0104)	(0.0291)	(0.0244)	(0.0154)	(0.0201)	(0.0150)	(0.0209)
-12	0.00951	0.0247	0.0349	0.0697	0.00817	0.0223	0.0174	0.0276
	(0.0138)	(0.0202)	(0.0343)	(0.0435)	(0.0165)	(0.0238)	(0.0178)	(0.0231)
-11	0.0144	0.00729	0.0192	0.00895	0.0129	0.00233	0.00823	0.000648

Continued on next page

Table 3.4.7. Impact of Credit Insurance on Average Funding Probability (with Borroower Information) – Continued

Week	Per listed loan				Per listed PP loan			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(0.0150)	(0.0131)	(0.0278)	(0.0271)	(0.0177)	(0.0160)	(0.0169)	(0.0165)
-10	-0.0117	-0.0139*	-0.0265	-0.0337	-0.0202**	-0.024***	-0.0186*	-0.0215**
	(0.00911)	(0.00737)	(0.0277)	(0.0236)	(0.00876)	(0.00721)	(0.0109)	(0.0103)
-9	0.0125	0.0130	0.0118	0.0110	0.00324	0.00724	-0.000775	-0.00303
	(0.0157)	(0.0165)	(0.0310)	(0.0330)	(0.0163)	(0.0193)	(0.0136)	(0.0139)
-8	-0.00752	0.00613	-0.0219	0.00597	-0.00281	-0.00132	-0.00620	-0.00652
	(0.00949)	(0.0189)	(0.0238)	(0.0419)	(0.0176)	(0.0200)	(0.0144)	(0.0148)
-7	-0.00276	0.00385	-0.0154	-0.00157	-0.00236	0.00604	0.000796	0.0121
	(0.0103)	(0.0138)	(0.0207)	(0.0254)	(0.0149)	(0.0209)	(0.0141)	(0.0189)
-6	0.0222	0.0232	0.0167	0.0287	0.0154	0.0155	0.0217	0.0229
	(0.0186)	(0.0190)	(0.0364)	(0.0393)	(0.0188)	(0.0207)	(0.0189)	(0.0210)
-5	0.0273	0.00777	0.0551	0.0191	0.0245	0.0134	0.0196	0.0128
	(0.0188)	(0.0125)	(0.0340)	(0.0268)	(0.0210)	(0.0202)	(0.0170)	(0.0179)
-4	0.00278	0.00562	0.00186	0.00830	0.00163	-0.00219	0.00210	-0.00360
	(0.0133)	(0.0146)	(0.0298)	(0.0322)	(0.0168)	(0.0162)	(0.0151)	(0.0143)
-3	-0.00185	0.00807	0.00351	0.0303	-0.000161	0.00191	0.00553	0.00625
	(0.0109)	(0.0157)	(0.0343)	(0.0421)	(0.0160)	(0.0181)	(0.0157)	(0.0172)
-2	0.0208	0.0405	0.0359	0.0770	0.0182	0.0304	0.0219	0.0294
	(0.0178)	(0.0254)	(0.0382)	(0.0475)	(0.0198)	(0.0262)	(0.0212)	(0.0268)
-1	0.00625	0.0110	0.0133	0.0211	-0.000374	-0.000410	-0.000210	-0.000698
	(0.0148)	(0.0167)	(0.0326)	(0.0356)	(0.0158)	(0.0165)	(0.0152)	(0.0163)
0	0.256***	0.165***	0.311***	0.206***	-0.00157	-0.00507	-0.000398	-6.32e-05
	(0.0290)	(0.0241)	(0.0300)	(0.0302)	(0.0153)	(0.0155)	(0.0142)	(0.0156)
1	0.539***	0.552***	0.485***	0.493***	0.0309	0.0312	0.0298	0.0294
	(0.0250)	(0.0285)	(0.0232)	(0.0261)	(0.0204)	(0.0227)	(0.0193)	(0.0222)
2	0.705***	0.690***	0.614***	0.607***	0.0113	0.0113	0.0115	0.00853
	(0.0183)	(0.0249)	(0.0191)	(0.0240)	(0.0177)	(0.0199)	(0.0170)	(0.0182)
3	0.809***	0.781***	0.700***	0.674***	0.00456	-0.000754	0.0147	0.0111
	(0.0122)	(0.0174)	(0.0157)	(0.0197)	(0.0167)	(0.0163)	(0.0192)	(0.0197)
4	0.804***	0.792***	0.683***	0.669***	0.0105	0.00963	0.0115	0.0136
	(0.0114)	(0.0150)	(0.0156)	(0.0186)	(0.0177)	(0.0191)	(0.0170)	(0.0188)
5	0.719***	0.783***	0.593***	0.644***	0.0811***	0.0679**	0.0611***	0.0469**
	(0.0168)	(0.0179)	(0.0188)	(0.0212)	(0.0311)	(0.0321)	(0.0223)	(0.0238)
6	0.742***	0.745***	0.638***	0.645***	-0.0148	-0.0157	-0.0130	-0.0127
	(0.0147)	(0.0184)	(0.0177)	(0.0205)	(0.0140)	(0.0152)	(0.0166)	(0.0190)
7	0.735***	0.711***	0.623***	0.606***	-0.0109	-0.00994	-0.0143	-0.0134
	(0.0139)	(0.0185)	(0.0172)	(0.0203)	(0.0127)	(0.0151)	(0.0115)	(0.0137)
8	0.766***	0.744***	0.650***	0.632***	0.00524	0.0107	0.00718	0.0102
	(0.0125)	(0.0166)	(0.0165)	(0.0196)	(0.0173)	(0.0219)	(0.0168)	(0.0198)

Continued on next page

Table 3.4.7. Impact of Credit Insurance on Average Funding Probability (with Borroower Information) – Continued

Week	Per listed loan				Per listed PP loan			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
9	0.776***	0.812***	0.678***	0.705***	-0.00393	-0.00456	-0.00400	-0.00668
	(0.0125)	(0.0137)	(0.0165)	(0.0181)	(0.0168)	(0.0178)	(0.0162)	(0.0158)
10	0.816***	0.824***	0.718***	0.725***	-0.00570	-0.00299	-0.0143	-0.0143
	(0.0117)	(0.0133)	(0.0159)	(0.0175)	(0.0223)	(0.0278)	(0.0165)	(0.0188)
11	0.799***	0.797***	0.693***	0.686***	0.0492	0.0517	0.0446	0.0540
	(0.0128)	(0.0162)	(0.0168)	(0.0195)	(0.0315)	(0.0338)	(0.0307)	(0.0360)
12	0.785***	0.755***	0.677***	0.649***	0.00633	0.000529	0.0127	0.00827
	(0.0132)	(0.0173)	(0.0172)	(0.0197)	(0.0236)	(0.0236)	(0.0255)	(0.0278)
13	0.757***	0.793***	0.666***	0.701***	0.0332	0.0431	0.0222	0.0308
	(0.0140)	(0.0142)	(0.0172)	(0.0183)	(0.0284)	(0.0316)	(0.0250)	(0.0293)
14	0.806***	0.840***	0.711***	0.738***	0.0103	0.00150	0.00242	-0.00711
	(0.0117)	(0.0123)	(0.0164)	(0.0179)	(0.0251)	(0.0213)	(0.0191)	(0.0162)
15	0.841***	0.829***	0.743***	0.733***	0.0877**	0.148**	0.0768**	0.129**
	(0.00960)	(0.0123)	(0.0148)	(0.0166)	(0.0391)	(0.0629)	(0.0334)	(0.0548)
16	0.843***	0.810***	0.737***	0.701***	0.0320	0.0187	0.0190	0.00921
	(0.00932)	(0.0134)	(0.0151)	(0.0184)	(0.0285)	(0.0255)	(0.0255)	(0.0233)
17	0.798***	0.791***	0.684***	0.679***	0.0250	0.0390	0.0214	0.0401
	(0.0117)	(0.0140)	(0.0172)	(0.0189)	(0.0249)	(0.0335)	(0.0202)	(0.0280)
18	0.753***	0.815***	0.653***	0.707***	0.0284	0.0286	0.0198	0.0212
	(0.0140)	(0.0136)	(0.0180)	(0.0187)	(0.0302)	(0.0327)	(0.0243)	(0.0277)
19	0.798***	0.796***	0.703***	0.704***	0.109***	0.0950**	0.0852***	0.0754**
	(0.0117)	(0.0141)	(0.0162)	(0.0182)	(0.0369)	(0.0384)	(0.0302)	(0.0328)
20	0.812***	0.786***	0.708***	0.687***	-	-	-	-
	(0.0110)	(0.0142)	(0.0164)	(0.0185)				
21	0.798***	0.767***	0.685***	0.651***	0.0146	0.00792	0.0133	0.00354
	(0.0120)	(0.0158)	(0.0172)	(0.0202)	(0.0239)	(0.0243)	(0.0212)	(0.0203)
22	0.701***	0.738***	0.620***	0.653***	0.0200	0.0381	0.0212	0.0286
	(0.0167)	(0.0184)	(0.0196)	(0.0212)	(0.0267)	(0.0382)	(0.0250)	(0.0330)
23	0.742***	0.749***	0.659***	0.660***	0.0283	0.0285	0.0228	0.0362
	(0.0140)	(0.0169)	(0.0177)	(0.0200)	(0.0274)	(0.0313)	(0.0238)	(0.0296)
24	0.774***	0.746***	0.686***	0.664***	0.0180	0.0117	0.0169	0.0117
	(0.0136)	(0.0183)	(0.0179)	(0.0209)	(0.0261)	(0.0254)	(0.0230)	(0.0237)
25	0.767***	0.704***	0.681***	0.638***	0.0464	0.0446	0.0372	0.0407
	(0.0133)	(0.0187)	(0.0176)	(0.0206)	(0.0287)	(0.0293)	(0.0250)	(0.0269)
26	0.697***	0.759***	0.608***	0.654***	-0.00192	0.00749	-0.00386	0.00202
	(0.0170)	(0.0181)	(0.0203)	(0.0219)	(0.0183)	(0.0282)	(0.0163)	(0.0233)
27	0.745***	0.773***	0.641***	0.652***	0.0199	0.00668	0.0147	0.00117
	(0.0156)	(0.0169)	(0.0198)	(0.0216)	(0.0232)	(0.0193)	(0.0188)	(0.0157)
28	0.782***	0.786***	0.676***	0.677***	0.0409	0.0692*	0.0206	0.0481

Continued on next page

Table 3.4.7. Impact of Credit Insurance on Average Funding Probability (with Borroower Information) – Continued

Week	Per listed loan				Per listed PP loan			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(0.0147)	(0.0175)	(0.0194)	(0.0218)	(0.0292)	(0.0404)	(0.0204)	(0.0312)
29	0.817***	0.784***	0.710***	0.682***	-0.00568	-0.00947	-0.00907	-0.0138
	(0.0116)	(0.0157)	(0.0175)	(0.0207)	(0.0166)	(0.0162)	(0.0147)	(0.0138)
30	0.783***	0.807***	0.672***	0.693***	-0.0167	-0.0127	-0.0189*	-0.0144
	(0.0136)	(0.0146)	(0.0189)	(0.0206)	(0.0124)	(0.0188)	(0.0115)	(0.0172)
Obs	39,401	39,249	39,228	39,079	19,135	18,146	18,284	17,343
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time-trend	Linear	Stoch	Linear	Stoch	Linear	Stoch	Linear	Stoch
BRW Info	No	No	Yes	Yes	No	No	Yes	Yes
Joint-test:								
pre	42.73	48	35.31	40.1	31.13	33.32	28.03	32.76
	[.062]	[.02]	[.231]	[.103]	[.409]	[.309]	[.569]	[.333]
post	1637.65	1512.45	1197.56	1059.69	57.07	49.09	49.72	45.99
	[0]	[0]	[0]	[0]	[.001]	[.011]	[.01]	[.024]

§ Borrower information includes age, marriage status, education, income, job type, work experience, company size, house and car ownership, mortgage and auto loan.

† Robust standard erros are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

‡ Joint-test: pre (post) tests jointly the 30 weekly effects before (since) Ucredit office opening. P-values are in brackets.

References

Allen, Franklin, Jun Qian, and Meijun Qian. 2005. "Law, Finance, and Economic Growth in China." *Journal of Financial Economics*, 77(1): 57–116.

Ashcraft, Adam B., and Joao A.C. Santos. 2009. "Has the CDS Market Lowered the Cost of Corporate Debt?" *Journal of Monetary Economics*, 56: 514–523.

Ayyagari, Meghana, Asli Demirguc-Kunt, and Vojislav Maksimovic. 2010. "Formal versus Informal Finance: Evidence from China." *Review of Financial Studies*, 23(8): 3048–3097.

Bailey, Warren, Wei Huang, and Zhishu Yang. 2011. "Bank Loans with Chinese Characteristics: Some Evidence on Inside Debt in a State-Controlled Banking System." *Journal of Financial and Quantitative Analysis*, 46(6): 1795–1830.

Bai, Ying, Yan Shen, Minggao Shen, and Zhong Xu. 2009. "Bank Size and Small- and Medium-sized Enterprise (SME) Lending: Evidence from China." *World Development*, 37(4): 800–811.

Bolton, Patrick, and Martin Oehmke. 2011. "Credit Default Swaps and the Empty Creditor Problem." *The Review of Financial Studies*, 24(8): 2617–2655.

Brandt, Loren, and Hongbin Li. 2003. "Bank Discrimination in Transition Economies: Ideology, Information, or Incentive?" *Journal of Comparative Economics*, 31(3): 387–413.

Cowling, Marc. 2010. "The Role of Loan Guarantee Schemes in Alleviating Credit Rationing in the UK." *Journal of Financial Stability*, 6(1): 36–44.

Cull, Robert, and Linxin Colin Xu. 2003. "Who Gets Credit? The Behavior of Bureaucrats and State Bank in Allocating Credit to Chinese State-Owned Enterprises." *Journal of Development Economics*, 71(2): 533–559.

Ge, Ying, and Jiaping Qiu. 2007. "Financial Development, Bank Discrimination and Trade Credit." *Journal of Banking and Finance*, 31(2): 513–530.

Gregory, Neil, and Stoyan Tenev. 2001. "The Financing of Private Enterprise in China." *Finance and Development*, 38: 14–17.

Hellmann, Thomas F., Kevin C. Murdock, and Joseph E. Stiglitz. 2000. "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" *The American Economic Review*, 90(1): 147–165.

Hirtle, Beverly. 2009. "Credit Derivatives and Bank Credit Supply." *Journal of Financial Intermediation*, 18: 125–150.

Huang, Yiping, and Xun Wang. 2011. "Does Financial Repression Inhibit or Facilitate Economic Growth? A Case Study of Chinese Reform Experience." *Oxford Bulletin of Economics and Statistics*, 73(6): 833–855.

Iyer, Rajkamal, Asim Ijaz Khwaja, Erzo F.P. Luttmer, and Kelly Shue. 2009. "Screening in New Credit Markets: Can Individual Lenders Infer Borrower Creditworthiness in Peer-to-Peer Lending?" NBER working paper 15242.

Jarrow, Robert A. 2011. "The Economics of Credit Default Swaps." *Annual Review of Financial Economics*, 3: 235–257.

Jensen, Robert. 2007. "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector." *The Quarterly Journal of Economics*, 122(3): 879–924.

Lardy, Nicholas. 2008. "Financial Repression in China." Peter G. Peterson Institute for International Economics working paper PB08-8.

Lin, Mingfeng, Nagpurnanand R. Prabhala, and Siva Viswanathan. 2013. "Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending." *Management Science*, 59(1): 17–35.

Linton, Katherine C. 2006. "Access to Capital in China: Competitive Conditions for Foreign and Domestic Firms." *Journal of International Commerce and Economics*.

Lu, Susan Feng, and Yang Yao. 2009. "The Effectiveness of Law, Financial Development, and Economic Growth in an Economy of Financial Repression: Evidence from China." *World Development*, 37(4): 763–777.

Manova, Kalina, Shang-Jin Wei, and Zhiwei Zhang. 2015. "Firm Exports and Multinational Activity Under Credit Constraints." *The Review of Economics and Statistics*, 97(3): 574–588.

McKinnon, Ronald I. 1973. *Money and Capital in Economic Development.* Washington, DC.:The Brookings Institution.

Michels, Jeremy. 2012. "Do Unverifiable Disclosures Matter? Evidence from Peer-to-Peer Lending." *The Accounting Review*, 87(4): 1385–1413.

Poncet, Sandra, Walter Steingress, and Hylke Vandenbussche. 2010. "Financial Constraints in China: Firm-level Evidence." *China Economic Reivew*, 21(3): 411–422.

Renton, Peter. 2012. "A Look back at Prosper 1.0: How Revelvant are the Numbers?" www.lendacademy.com report.

Riding, Allan, and George Haines, Jr. 2001. "Loan Guarantees: Costs of Default and Benefits to Small Firms." *Journal of Business Venturing*, 16(6): 595–612.

Saretto, Alessio, and Heather E. Tookes. 2013. "Corporate Leverage, Debt Maturity, and Credit Supply: The Role of Credit Default Swaps." *The Review of Financial Studies*, 26(5): 1190–1247.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing Like China." *The American Economic Review*, 101(1): 196–233.

Stiglitz, Joseph E. 1994. "The Role of the State in Financial Markets." 19–61, World Bank. Washington, DC.