Essays on Peer-to-Peer Lending

Jin Huang

A thesis submitted for the degree of Doctor of Philosophy

Essex Business School

University of Essex

March, 2021

Abstract

The Peer-to-Peer (P2P) lending model has become increasingly popular in China in recent years. In 2012, there are only 298 P2P platforms operating in China and loan volume is 22.9 billion RMB while in the first half of 2018, there are 1881 P2P platforms and trading volume has reached 7.33 trillion RMB. Although both number of platforms and transaction volume have increased significantly, severe asymmetric information still discourages participants. This doctoral thesis uses three empirical chapters to investigate the P2P lending market in China. Drawing on Message framing and signaling theory, we first examines the extent to which message framing is associated with funding outcomes receive in the context of P2P lending and whether positive message framing reinforces the positive impact of credit ratings on funding outcomes. Using a Heckman two stage model, we find that the use of positively framed messages is positively associated with positive funding outcomes. Besides, positive message framing enhances the positive impact of the credit ratings (an example of costly signals) on funding outcomes.

We then investigate the role of psychological distancing and language intensity in P2P funding performance. We bridge the P2P lending literature and psycholinguistics literature and set out to explain how psychological distancing manifested by linguistic styles can influence lenders' decision on P2P funding campaign. We find that linguistic styles related to psychological distancing are negatively related to P2P funding success. Moreover, the language intensity tends to strengthen the negative relationship between psychological distancing and funding success. This finding is consistent with psycholinguistics literature which suggests that psychological distancing is associated with negative interpersonal outcome (Simmons et al, 2005; Revenstorf et al, 1984). Specifically, the number of "you" and the number of negations used in borrowers' description are negatively related to the willingness of the lender to support the funding campaign. The intensive language negatively strengths the relationship between the funding performance and number of "you" but does not apply to number of negations.

Lastly, we investigate the funding performance of the financial excluded borrower in a large P2P lending platform. The association of financial technology (fintech) and financial exclusion has attracted attention since rapid growth of fintech innovation. Using loan-level data from a lending Chinese P2P company, we find there is a negative indirect effect of financial exclusion on funding success through credit score. In a moderated mediation analysis, we also find new business model such as offline authentication and education qualification positively moderates the linkage between the financial excluded and credit score and therefore negative indirect effect of financial exclusion on funding success is overturned when the excluded borrower has conducted offline authentication and obtained higher education qualification. In the end, we examine the determinants of offline authentication decision. We find the borrowers in a city with better financial infrastructure are more willing to conduct authentication. However, the financial excluded borrowers are less likely to conduct offline authentication.

Keywords: P2P Lending; Message Framing; Psychological Distancing; Financial Exclusion

Acknowledgements

I started the PhD journey in 2017 now it is coming to an end. Looking back, the PhD journey is never an easy path. Without generous help from my supervisors, parents, and friends, I cannot accomplish this achievement. I would like to take this opportunity to express my sincere gratitude to those who offered me selfless support and walked along this journey with me.

First and foremost, I would like to express my highest gratitude to my first supervisor Dr Jun Li for dedicatedly and patiently leading me through each stage of my PhD study. Dr Jun Li makes substantial contributions to my PhD thesis. Without his commitments, thesis would not be the same. It is from Dr Jun Li that I learnt how to conduct high quality academic research. The knowledge he delivered to me will continuously influence my future career. I see Dr Li not only as a committed supervisor but also a reliable friend who have always supported me in both good and bad times in my life.

I am grateful to my second supervisor Prof Vania Sena. She is my supervisor since the time I entered to the PhD program at University of Essex. Although Prof Sena has a lot administrative commitments, she has followed my progress very closely. She offered me selfless support, and inspired me to become a competent researcher.

Undertaking PhD research sometimes is boring. However, I am lucky to have some friends who can always bring funs to me. My special thanks go to those friends in a Wechat group called “Colchester Midnight Poker”. I will always remember the nights we played poker and hung out together.

Finally, I am heavily indebted to my parents, for their continuing support and unconditional love. Their love and support are my source of strength that helped me overcome difficulties.

Table of Contents

List of table

List of figures

Chapter 1 Introduction

1.1 Research Background

Peer-to-peer lending, abbreviated P2P lending, is the transaction of lending money through online services to individuals or businesses. Scholars define P2P lending as unsecured lending relationship that cuts off financial intermediary between borrowers and lenders via online platform (Lin et al, 2009; Bachmann et al, 2011; Tao et al, 2017). Since, at beginning, the P2P lending platforms offer these services entirely online, they can operate with lower overhead and keep the service in a lower price than traditional financial institutions. Therefore, through online transaction, P2P lending makes borrowers without going through any traditional financial intermediaries (Berger and Gleisner 2009; Wang et al. 2009; Bachmann et al. 2011).

Since the launch of the first Peer-to-Peer (P2P) lending platform, Zopa, in 2005, the P2P lending has become increasingly important in credit market for both consumers and entrepreneurs. It is part of a larger crowd funding movement which adopts internet to get access to collective funding (Burtch et al. 2013). This lending marketplace is created to supplement traditional bank lending to meet the needs of individuals and Small and Medium Enterprises (SMEs). For example, P2P lending helps the people who find difficult to obtain loans from financial intermediaries such as banks get access to the funding they need.

The Chinese P2P market begins at 2007 since the launch of the first platform, Paipaidai. During recent years, the P2P lending market has grown very fast in China because of growing credit demand from both entrepreneurs and consumes. According to the data provided by wdzj.com, a third party website that provides a comprehensive data for Chinese P2P lending market, there are 1881 P2P platforms operating up to the first half of 2018. The total accumulative transaction volume of P2P lending platforms has reached 7.33 trillion RMB. In the first half of 2018 the number of online borrowers and investors respectively reached 4.35 million and 4.08 million. However, in 2012, there are only 298 P2P platforms and loan volume is 22.9 billion RMB.

The marvelous growth of P2P industry can be attributed to following reasons (Xu, 2017). First, the financial services are insufficient from both demand and supply side. On the demand side, households who have financial needs and small and medium enterprises (SMEs) are underserved by existing financial system. The big and connected companies such as State-Owned enterprises (SOEs) have access to relatively sufficient funds at reasonable cost. Yet, SMEs do not get funding opportunities or pay high price. Chinese financial institutions prefer large firms and the financial system is designed for their needs.). On the contrary, the financial institutions in the US provide services to all types of customers including small business and households. There are a large number of community and regional banks to serve small firms and households. The Chinese government therefore adopts a relatively tolerant attitude toward fintech innovation aiming to provide financial services to a large group of people (Wang et al., 2016). Taking P2P lending for example, it was not clear which regulatory authority should oversee P2P platforms for many years and no special license is required to run the P2P platform (Wang et al., 2016). Firms from other industries can also easily expand their business into fintech industry. For example, the Chinese government issued financial licenses to many technology companies such as Alibaba, Jingdong and Tecent enabling them become fintech firms. Even in the developed countries, it is impossible to see a highly regulated industry to accept such firms. On the supply side, wealth and middle class households find difficult to make good investment. Over the past decades, the income growth and saving rate have increased significantly, which creates tremendous demand for wealth management. They may eventually become active investment opportunity hunters in the P2P market.

Second, the advances in information technology lift the barriers of financial services, expand the customer base and create the new business model. For example, the big data collection and analysis enable financial institutions to deliver their services to larger group of people. The technology can be more accurate in determining creditworthiness of borrowers who do not have credit history and extend credit to them. Third, the loose regulation in the industry leads to abnormal business adventure and risk-taking including Ponzi schemes. There are no proper regulations for the industry before 2015. Only after excessive frauds were unveiled, the initial regulations were proposed in 2015 and came effective in 1st July 2016. The tolerant attitude to P2P lending can be attributed to the desire of developing better financial system. The notion that China's financial system is underdeveloped is well-accepted. Therefore, anything that can contribute to the development of the system is encouraged but sometimes without prudent supervision. Many platforms take risks and conduct Ponzi-like operation. They collect funds from new customers and pay interest to the old ones. They use rest of the money to invest in highly risky projects and hope to make swift gain. If they made wrong investment decision, they just lose other people's money. This is an asymmetric information that probably persists in all financial institutions. Due to the tolerant attitude of the authority, it is not uncommon that Ponzi-like platform flourished, which results in small-scale crisis.

Although P2P lending platform is an information media between borrowers and lenders, the internet-mediated environment not only produces useful information but noise and spurious information as well. To get the project funded, borrowers will always pretend to have good credit history and projects. Without investment experiences and deep investigation, it is difficult to identify good borrowers from bad borrowers. Moreover, it is not cost efficient for individual investors to carry out information analysis. Traditionally, specialized financial institutions are responsible for solving the information asymmetry. Yet, asymmetric information online contains much more noise and it is difficult for participant to differentiate.

The difficulties that P2P industry is experiencing are largely due to information asymmetry among lenders, borrowers and the platform. Traditional financial institutions reduce asymmetric information between lenders and borrowers in two ways, relationship approach (by building long term relationships with customers and collecting soft information) and the transaction approach (which relies on the collection of hard information through financial statement and historical credit data) (Berger and Udell, 2002). Yet, in P2P industry the relationship approach cannot be used as the platform is structured in such a way that long-terms relationships cannot be built. As for the second approach, platforms may not have operational for a period of time long enough to be able to collect sufficient hard information on the borrowers.

Asymmetric information has been noticed in P2P lending literature. Researchers have suggested that asymmetric information problem can be alleviated if borrowers can send costly signals to lenders (Caldieraro et al., 2018). Lin et al (2013) finds that social capital increases P2P funding success, lowers interest rate on funded loans and decrease the default rate. Liu et al (2015) find when offline friends place bid in P2P lending campaign, potential investors are willing to follow with a bid. However, if lenders lack of investment knowledge and experiences, which is common in the P2P lending market, checking costly signals is challenging (Anglin et al, 2008). Moreover, costly signals are not easy to obtain in the P2P market (Loewenstein et al., 2014). Therefore, the first study (chapter 2) suggests borrowers' message framing that communicates their trustworthiness to potential lenders is a solution of asymmetric information. Drawing upon message framing theory and costless signaling theory, the paper arguestrust-related message written in borrowers' loan application positively predicts the probability of funding success and the message framing can also reinforce the effect of credit rating on funding success.

The results from first study are consist with existing literature which shows language plays an important role in crowdfunding (Herzenstein et al, 2011; Dorfleitner et al, 2016; Defazio et al., 2020; Huang et al., 2020; Parhankangas and Renko, 2017; Larrimore et al., 2011). However, not only what they said matters but how they conveyed message influences likelihood of funding success as well. The questions of "how" is related to linguistic styles. Unlike content words such as adjectives, verbs and nouns, style words do not carry much meanings. Yet, psycholinguistics literature suggest style words are associated with individuals' social and psychological worlds and contribute to appeal to audiences for communicators (Pennebaker and Chung, 2013; Tausczik and Pennebaker, 2010; Parhankangas and Renko, 2017). Extant research shows that the frequency of certain style words we use is associated with how we are perceived by others and has an impact on outcomes such as academic performance and funding campaign of social entrepreneurs (Robinson et al.,2013; Parhankangas and Renko, 2017; Fausey and Boroditsky, 2010). Furthermore, Larrimore et al (2011) find stylistic features of messages can be effective in assessment of source trustworthiness especially in an online environment where social cues such nod and smile about the information source are less reachable. Therefore, it is expected that P2P lenders are also sensitive to the use of linguistic styles of borrowers.

Certain style words are closely related to psychological distancing. Parhankangas and Renko (2017) define Psychological distancing is to the extent to which people distance or remove themselves away from the topic being discussed. The borrowers may unconsciously use of certain style words that are associated with psychological distancing and disconnect to the lenders on an emotional level because certain language use can distance away relationships (Tausczik and Pennebaker, 2010). Similarly, Nook et al (2017) suggests that psychological distance is embedded in language use. For example, use of present-tense verbs and first person pronouns which are commonly used style words imply temporal and social closeness rather than distancing (Mehl et al., 2013; Pennebaker and King, 1999). Hence, the second study (chapter 3) argues that psychological distancing measured by style words alienates the P2P borrowers from their crowds. Given the tone is an important determinants of communication quality, it also suggests that intensive tone negatively moderates psychological distancing and funding performance.

The purpose of establishment of P2P platform is to meet the financial needs of the borrowers. P2P platforms therefore are similar to microfinance institutions in many ways although P2P platforms mainly operate online. One of the main objectives of them are allaying financial exclusion. Kempson and Whyley (1999a,b) have identified five dimensions of financial exclusion including access exclusion, condition exclusion, price exclusion, marketing exclusion and self-exclusion. The third study (chapter 4) focuses on difficulty of accessing financial resources and it is a supply-side issue of financial exclusion. To alleviate financial exclusion is a critical element for poverty reduction and economic growth. For example, using a natural experiment in Mexico, Bruhn and Love (2014) find more access to financial services results in an increase in income for low-income people and a decrease in unemployment. Due to only a few people who are able to access to basic financial services nowadays, alleviating financial exclusion is yet an unfinished task.

P2P lending offers a potential avenue for addressing the financial exclusion. Komarova Loureiro and Gonzalez (2015) suggest due to substantial bureaucracy and paperwork in traditional financial institutions, it is costly to get access to funding while the recent technology has overcome this difficulty and can complete the deal within few days by P2P platforms. However, the empirical evidence regarding P2P lending and financial exclusion is in general mixed (Komarova Loureiro and Gonzalez., 2015; Lin and Viswanathan, 2015). The results from third study (chapter 4) shows that P2P lending doesn't necessarily alleviate financial exclusion. Due adverse credit score, the financial excluded are less likely to be funded. In addition, it finds through platform offline authentication, the financial excluded can increase their chance of getting funds. It also shows that the excluded borrowers will have a better chance if they obtained a higher education qualification.

Asymmetric information and P2P lending in China

The asymmetric information means one party has better information than the other party in the transaction. This asymmetry among the parties may result in a failure in the functioning of the market (See Akerlof, 1970). For example, in the P2P lending market, all the entrepreneurs claim that they have promising projects and are worth for the fund. Yet, it is probably that only half are "good". How to identify the "good" has attracted attentions by the related parties. Spence (1973) posits that the asymmetric information can be alleviated when the high-quality entrepreneurs send "signals" to communicate their quality. Sending signals is costly so low quality entrepreneurs are unable to send these "signals". In the traditional credit market, the lenders or investors can require collaterals provided by the borrowers to signal their creditworthiness. However, in P2P market, collaterals are normally not required and therefore it is more difficult to signal creditworthiness. Actually, P2P lending is mainly unsecured loans. These loans are provided to those who have difficulties to acquire loans from traditional financial institutions such as banks (Milne and Parboteeah, 2016). P2P platforms are proud of that they can serve a larger group of costumer at a relatively low cost. Yet, Freedman and Jin (2011) argue compared to offline credit markets, P2P lenders encounter severer asymmetric information because people cannot authenticate the online information.

Signaling theory has been therefore widely applied to crowdfunding literature (Anglin et al, 2018; Steigenberger and Wilhelm, 2018; Kromidha and Robson, 2016; Ahlers et al.,2015).Traditional signaling theory pays closer attention to costly signals. For instance, good education background of job candidates can serve as a meaningful signal that they are skilled workers to employers (Spences, 1973). Since obtaining a degree from good university is costly, candidates with this degree are highly likely to be good quality. In the context of crowdfunding, costly signals such as use of media, human capital, social capital and intellectual capital have an influence on crowding performance (e.g., Anglin et al, 2018; Ahlers et al.,2015; Courtney et al, 2017). In addition, the theory concludes that language-based signals are in general inefficient because they are costless for both high and low-quality senders (Steigenberger and Wilhelm, 2018; Farrell and Rabin, 1996). Receivers are also more likely to rely on costly signals that are able to separate high and low quality to make decisions. Moreover, Farrell and Rabin (1996) refer to language-based signals as cheap talk.

Yet, management studies and practices emphasize the importance of costless signals. Researchers find that language that consists of charisma, confidence, optimism, resolve, narcissism and entrepreneurial passion has a great impact on resource acquisitions (Martens et al, 2007; Aktas et al, 2016; Davis et al, 2017; Avey et al., 2011). For example, Martens et al (2007) identify several identities such as aspiration and ambition within IPO prospectuses significantly influence IPO process. Scholars investigate in the role of language in P2P market as well (e.g., Herzenstein et al, 2011; Michels, 2012; Dorfleitner et al, 2016). Herzenstein et al (2011) examine whether identity claims invoked in borrowers' language influence lender decision using Prosper dataset and they find larger number of identity claims in borrower description will increase funding success rate and lower the interest rate changed but a negative association between the number of identities and loan performance. Dorfleitner et al (2016) studying two German platforms find spelling errors, text length and positive emotion evoking keywords have an impact on funding success but little influence on delinquency. Consistent with Dorfleitner et al (2016), using a Chinese P2P dataset, Han et al (2018) find completeness of description and positive sentiment language are more likely to get funded. In addition, the communication between lender and borrower is able to result in funding success. The number of borrowers' responses positively affect funding success but lenders' comments influence funding success negatively (Xu and Chau, 2018).

Asymmetric information is more severe in Chinese P2P market because Chinese P2P lending platform uses their own credit rating system to evaluate the trustworthiness of borrowers, unlike UK or US where the credit score is assigned by specialized and independent rating agency (Tao et al, 2017). There is lack of reliable personal credit rating agencies in China, although Ant financial (a financial subsidiary of Alibaba group) provides credit scores to individuals. The scores assessed by Ant financial are biased because they are likely to give good rating to the people who use the services provided by Alibaba. So, Chinese platforms nowadays are not willing to adopt Ant financial's credit rating and assess the credit history of the borrowers by themselves. Moreover, the majority of Chinese P2P platforms don't apply Web 2.0 functionality which enables communication between lenders and borrowers. This means that most of Chinese P2P investors can only see static listing information and are not able to directly exchange information with the borrowers while investors in Prosper (a US based P2P platform) can communicate with the people who would like to support. The only exception is that Xu and Chau (2018) find a unique dataset from a Chinese P2P platform and examine the impact of lender-borrower communication on funding success but they use a fictitious company name for confidentiality. In addition to this, another key difference is that unlike developed economies that use a pure online lending process, Chinese P2P platform provides offline authentication service to further mitigate asymmetric information that is originated by pure online authentication. Despite the important role of offline authentication, the study on offline authentication is scarce. This thesis tries to fill this gap by investigating the role of offline authentication in different context.

1.2 Research motivations and research questions

The credit risk management in the traditional credit market is mature. The borrowers' collateral and credit score will determine the loan terms such as the loan amount and interest rate. The financial institutions will conduct periodic review to ensure the borrower will pay back the loan. However, the P2P industry is lacking such process and the credit risk therefore cannot be well controlled. Due to the imperfection of the online lending process, borrowers strategically hide their credit information to obtain funds from the market. Funds are limited resources. When the borrowers cannot pay back the loan, the total amount of money will be less and new investors are hesitant to enter the market. If so, financial inclusion which is one of the main objectives of financial technology cannot be achieved. Hence, an important question in P2P market is how participants mitigate risks associated with information asymmetry. To better understand this question will also be beneficial to participation in the market and financial inclusion.

Studies have suggested that language and associated message framing (two cheap signals) are able to address the risks generated by information asymmetry as they would help lenders to differentiate between high and low-quality borrowers (Anglin et al., 2018; Loewenstein et al., 2014). Cheap signals are useful especially when costly signals are difficult to obtain and people are lack of experience to judge costly signal (Loewenstein et al., 2014; Anglin et al, 2008). Yet, extant research finds mixed evidences regarding cheap signals in online environment. For example, Ludwig et at (2013) suggest message framing does not perform well in online review evaluation while Anglin et al (2018) shows cheap signals that indicate a successful borrower have a positive effect on crowdfunding performance. Existing studies examining the role of cheap signals in the P2P lending have focused on a small set of signals such as photographs, message length, text descriptions and spelling mistakes (Dorfleitner et al., 2016; Duarte et al., 2012). These papers provide limited insights into the impact of message framing on a lender's decision in the context of P2P lending. Therefore, the study 1 asks: Does cheap signals associated with language have an impact on funding performance? Does cheap signals complement traditional costly signals in the evaluation of the trustworthiness?

Moreover, studies suggest borrowers' language normally focus on different aspects in different types of funding campaigns (Parhankangas and Renko, 2017; Anglin et al., 2018; Block et al., 2018; Herzenstein et al., 2011; Allison et al., 2013; Majumdar and Bose, 2018) . In reward-based crowdfunding, the language often centers on new product development plan, while in P2P lending, borrowers tend to stress on evidence-based elements because personal details and loan characteristics can appeal more lenders (Lee et al., 2019; Allison et al., 2013). Extant studies have tested linguistic styles, measured by use of some style words and found lenders form perceptions about prospective borrowers based on the style words and therefore style words influence lenders' decision (Parhankangas and Renko, 2017). Due to different focus of entrepreneurial narratives in P2P lending, it is important to investigate how P2P lenders form the perception of linguistic styles. Moreover, the language intensity can also affect the funding outcome (Han et al., 2018). To date, researchers have studied these two factors in isolation. Yet, linguistic styles and language intensity are actually inseparable. Lenders receive both together to form opinions about the borrower. In addition, scholars tend to pay attention to how positive linguistic styles improve likelihood of funding success. Extant literature largely ignores the negative effect of some linguistic styles such as psychological distancing on funding outcomes (Huang et al., 2020). Given the fact that the important role of psychological linguistic styles on online context (Ludwig et al., 2013; Peng et al., 2004), the study 2 asks: Does psychological distancing affects P2P funding outcomes and Does language intensity strengthen this relationship?

P2P lending enables borrowers to get unsecured loans (e.g., loans without collateral) from individual lenders in a P2P platform (Lin et al., 2013). The unsecured loans lift the barriers of people getting the loan to some extent so researchers consider P2P lending may allay financial exclusion (Sparreboom and Duflos, 2012). The study to examine the interplay between financial exclusion and P2P lending success is scarce and contradictory (Komarova Loureiro and Gonzalez., 2015; Lin and Viswanathan, 2015). More importantly, little is known about the role of credit score, offline authentication and educational attainment in the linkage between the financial exclusion and P2P funding success. The research gap is problematic given the asymmetric information of financial excluded borrowers can be reduced by these factors, thereby promoting participation from the lenders and contributing to financial exclusion. Related parties such as P2P platforms and the government can act accordingly based on these information. We first begin to fill the gap by examining the mediating role of credit score. Does the credit score play a mediating role in the association between the financial excluded and funding success? Second, we test moderated mediating of role of offline authentication and human capital. Does offline authentication and human capital positively moderate the financial excluded-credit score-funding success linkage?

1.3 Data and Methodology

1.3.1 Data

The P2P platform in China offers us a laboratory to study the questions mentioned above. First, the number of P2P lending participants in China has increased significantly in recent years. Second, there are many P2P platforms in China providing offline authentication to mitigate asymmetric information given the asymmetric information is particularly severe in the market. The following three empirical studies will therefore mainly use the data from Renrendai. Renrendai is one of the largest P2P platforms operating in Mainland China and many papers use Renrendai's data due to representativeness. (Mi and Zhu, 2017; Tao et al, 2017 and Yao et al., 2018). Renrendai was established in 2010 and registered capital is 100 million RMB. In January 2014, Renrendai acquired venture capital (130 million dollar) from TrustBridge Partners. Until Oct 2018, the accumulative transaction reaches 71.4 Billion RMB (Renrendai.com). Now it has been ranked 2nd by wdzj.com among all 1881 P2P platforms in China and has more than 1 million members located in over 200 cities (Chen et al., 2019). The loan application process is very similar among Chinese P2P platforms although each platform has their own in-house credit rating system.

The loan application process is as following. Frist, borrowers submit their application form with their national ID number and other personal information. They would need to specify the requested loan amount, the interest rate they would like to pay, the duration they will pay back the loan, the purpose of borrowing and any other information they find helpful to their application. The platform will then access their ID, mobile number, address, employment, income and etc. it is suggested by the platform to disclose additional information such as education qualification, car and house ownership, marital status and other professional certification to promote their credit rating. At the end, the platform will verify the information submitted by applicants and assign a credit grade to the applicant. Each credit grade corresponds different credit rating from HR (high risk) to AA (very safe). Given the similar operation models of the P2P platforms in China, the Renrendai we select is not an outlier. We crawled the data between 1 Jan 2015 and 31 Dec 2015 from the Renrendai official website. The information includes loan information (e.g., interest rate, loan amount and etc), demographic information (e.g., borrowers' age, education attainment, gender and etc) and borrowers' narratives. The total number of listing is more than 400,000, which is consistent with prior studies (Chen et al., 2018).

1.3.2 Methodology

Heckman two stage model are applied in the study 1 and study 2. The adoption of Heckman two stage model is to address sample selection issue that widely exists in the literature. The scholars usually delete P2P loans with offline authentication when they study loan description (e.g., Chen et al., 2018). However, we argue that simply removing these listings will result in sample selection bias. We therefore apply Heckman model to solve the issue. The Heckman model has two step: the first step is to estimate the likelihood of borrowers choosing offline authentication by applying a probit model to whole sample including all listings with and without offline authentication. In the second step, the sample is restricted to the listings without offline authentication. Then, we estimate the effect of words written in borrowers' description on funding performance. As both offline authentication and funding success are dummy variables, we apply a probit with sample selection model, so called heckprobit. Although we call it two step, the software estimates first and second-step simultaneously by using a maximum likelihood estimation approach (Andres, 2014). Identification of selection equation requires at least one variable that influences the choice of offline authentication but not affect investors' decision (Pham and Talavera, 2018). Branch will be used to meet this exclusion restriction. Branch is dummy variable, taking value of one if there is an offline branch in borrower's city or zero otherwise. This variable is valid because lenders' decision is unlikely to be affected by a city's offline branch and the establishment of offline branch affects borrowers' offline authentication decision. In the second step, only the sample without offline authentication is selected. Control variables include some demographic and listings factors such as age, degree, income, interest rate, log loan amount, duration and credit rating.

The study 3 adopts structural equation modeling (SEM) to test the mediation /moderated mediation effect of credit score, offline authentication and human capital in the relationship between the financial excluded and funding success. There are many studies that recommend SEM approach to test mediation effect (Iacobucci et al., 2007; Zhao et al., 2010; Cho and Pucik., 2005). Using SEM instead of three separate regressions proposed by Baron and Kenny (1986) can better control for measurement errors which might result in under- or over-estimation of mediation effects (Shaver, 2005). In addition, it can also estimate everything simultaneously rather than assume the equations are independent. However, the nature of our dependent variable (funded is a dummy variable) makes a linear SEM ill-suited. Applying a linear model with dummy dependent variable will lead to biased results. Hence, we use Generalized SEM (GSEM) model that allows binary outcome to fit our proposed regressions (Kaplan and Vakili., 2014). We apply linear regression to first part of analysis (determinants of credit score) and logit regression with a dummy dependent variable to second part (determinants of funding success). Bootstrapping is used to estimate standard error and confidence intervals for the indirect effects.

1.4 Chapter conclusion

The introductory chapter shows an overview of following three empirical studies. It provides background information of P2P market, motivations and key research questions of the empirical chapters, and data sources as well as methodology for each study.

Chapter 2 Message Framing in P2P Lending Relationships

Abstract

The purpose of this paper is two-fold: first, it examines the extent to which message framing is associated to funding success and to the number of bids projects receive in the context of P2P lending; second, it investigates whether positive message framing reinforces the effect of credit rating on funding success. Our analysis is conducted on a dataset of 33028 listings of potential borrowers from a Chinese P2P lending platform using a Heckman selection model. We find that the use of positively framed messages through written language positively predicts the likelihood of funding success although it is not correlated to the number of bids the project receives. In addition, the use of positive language enhances the effect of borrowers' credit rating on funding success. Our results contribute to the literature on the effectiveness of cheap signals in the context of Internet-based interactions while highlighting complementarities between different types of signals in the context of P2P lending.

2.1 Introduction

Financial technologies (‘FinTechs’) are changing the face of global finance by integrating finance and technology in ways that disrupt traditional financial models while providing an array of new services to businesses and consumers. The hybridization of technology with the traditional processes of finance has replaced traditional structures as well as leveraging alternative business models that take advantage of the new technologies and market conditions. Importantly, Peer to Peer (P2P) lending (thereafter, P2P lending) has given rise to a new means of finance, where individual lenders make unsecured loans to other borrowers seeking funds for individual and/or business purposes (Lin et al., 2013).

Traditionally, financial institutions deal with asymmetric information on the quality of the borrowers either by building long-term relationships with customers so that “soft” information on loan applicants can be collected or by developing credit score systems that rely on the collection of “hard” information from loan applicants through transactional data (Landström, 2017). However, these two strategies to mitigate the risk posed by asymmetric information on the quality of the borrowers are not entirely viable in the context of the P2P platforms. Indeed, P2P platforms are structured in such a way that long-terms relationships with customers cannot be easily built while at the same time, platforms may not be in a position to easily collect “hard” information on their clients given the fact that they have not been operating for a long time. In other words, “centralized” solutions led by the platforms themselves are not viable.

In the case of P2P platforms, it has been argued that lenders’ message framing aiming at communicating their trustworthiness to potential borrowers can be a solution to the risks generated by asymmetric information. Message framing through language is usually considered to be a low-cost (Anglin et al., 2018) and therefore different from costly signals which are less imitable and more valuable and can facilitate the sorting between high quality signalers from low quality signalers (Anglin et al. 2018). For example, when lenders lack sufficient experience and knowledge about an investment, which can often be observed in P2P lending context, assessing costly signals such as credit rating of a borrower becomes challenging (Anglin et al, 2018).

Cheap signals[1] are particularly useful when costly signals are difficult to collect (Loewenstein et al., 2014). However, previous studies have found mixed results about the effectiveness of message framing, usually considered a cheap signal, in the context of Internet-based interactions. In particular, while Ludwig et al.'s (2013) study have shown that message framing may not always be important in the evaluation of online reviews, Anglin et al. (2018) observe the effectiveness of projecting the attributes indicative of successful lenders through language, as a cheap signal, on crowdfunding success. Studies examining the role of cheap signals in the context of P2P funding have focused on a small set of signals such as photographs, message length, text descriptions and spelling mistakes (Dorfleitner et al., 2016; Duarte et al., 2012). However, these papers provide limited insights into the impact of message framing on a lender's decision in the context of P2P lending. Still, this issue is important to investigate since the uncertainties of P2P lending decisions ex-ante and the associated costs of these decisions ex-post (Dorfleitner et al., 2016; Duarte et al., 2012; Guo et al, 2016) are higher than those involving other forms of internet-based interactions such as online reviews.

While positive framing of a message through written language can overcome the challenges of assessing the soft qualities of borrowers due to the limitations of forming long-term relationships in the context of P2P lending, it may also provide complementary knowledge in the evaluation of costly signals about borrowers such as their credit ratings. The study also suggests that message framing including cognitive and affective attributes such as trustworthiness can play a significant role in complementing costly signals such as credit ratings. The study contributes to previous research on P2P lending by investigating not only the influence of message framing through

[1] The concept is rooted in the signalling theory and refers to the fact that some signals individuals transmit can be either costly (i.e. a certain amount of time and effort is required to be able to send the signal) or cheap (i.e. very little effort is required to be able to send the signal).

language but also the complementary role of different signals in affecting funding decisions within P2P lending settings (Dorfleitner et al., 2016; Duarte et al., 2012; Guo et al, 2016; Davis and Allison, 2013).

Against this background, the purpose of the paper is two-fold. First, it examines to what extent message framing is correlated to funding success in the context of P2P lending using insights from the signaling and message-framing theories. Second, it assesses how it complements credit rating in conveying potential lenders soft information on the characteristics of the lenders so that the project may be funded. Our empirical analysis is conducted on a data-set of 33028 listings drawn from the leading Chinese P2P platform, Renrendai. The results suggest that the use of positive message framing in the description of the project positively predicts the likelihood of funding success but not the number of bids. In addition, message framing enhances the effect of borrowers' credit rating on funding success.

The structure of the paper is as follows. Sections 2 develops our set of hypotheses on message framing and funding success. The empirical methodology and the data are presented in Section 3. The empirical results are discussed in Section 4 while the robustness tests are presented in Section 5. Finally, a discussion on the implications of the results is presented in Section 6 while some concluding remarks are offered in Section 7.

2.2 Theoretical background

Signaling theory focuses on the mechanisms that may reduce asymmetric information between two parties. More specifically, it focuses on whether and how a more informed sender (e.g. the borrower) communicates information (i.e. a signal), and how a less informed receiver (e.g.

potential investor) interprets the signal (Lee et al., 2016). Communication between parties can be realized by sending costly signals which are based on observable actions or qualities and/or cheap signals which are less observable and rely mainly on words or cheap talk (Cheung et al., 2014). Traditionally, signaling theory argues that costly signals are more effective in solving the asymmetric information problem than cheap signals. For example, while the costs of obtaining IS09000 certification are high due to its labour-intensive and time-consuming nature, such costs improve communication of quality signals and make false signaling more difficult (Connelly et al., 2011). In this context, the traditional argument of the signaling theory suggests that evaluating quality from a cheap signal is challenging (Anglin et al, 2018; Farrell and Rabin, 1996; Steigenberger and Wilhelm, 2018) as there is little or no explicit cost to acquire and send, and therefore it is effective when objective information about the sender is unavailable. There are many examples this is the case. For instance, Chen et al. (2009) find that entrepreneurial passion (a cheap signal) is an intangible and hard-to-measure quality of potential entrepreneurs which however does not help them in acquiring resources through venture capital. Thus, venture capitalists tend to rely on costly signals about the quality of the entrepreneurial venture. Yet, these studies focus on very specific circumstances when cheap gnals can deliver valuable information to receivers (e.g. investors) in an effective way (Danilov and Sliwka, 2016; Marti and Balboa, 2007).

Cheap signals such as message framing are particularly useful when the audience lacks sophistication and the costly signals received are difficult to assess (Loewenstein et al., 2014). This can be relevant to the case of P2P platforms. Indeed, investors of P2P lending platforms are often challenged when assessing the costly signals associated with it (Anglin et al, 2018). In these cases, message framing may be one of the most important cheap signals that potential borrowers may have to convince lenders to support the project (Anglin et al, 2018; Steigenberger and Wilhelm, 2018). As such, signaling theory can be complemented by message-framing theory: it suggests that whether the way the message is framed through language will act as a signal which can affect a receiver's (i.e. investor's) actions or decision-making (Lee and van Dolen, 2015). The literature on message-framing mostly agrees that consumers with low level of involvement in a decision are strongly influenced by positively framed messages (Bester and Jere, 2012; Maheswaran and Meyers-Levy, 1990). In online environments, the use of language through textual messages and interactions for sharing sentimental or affective expressions, and cognitive messages is the main means of communication (Lee and van Dolen, 2015). As opposed to offline contexts in which verbal and facial communication is used, in online environments, users affectively and cognitively frame messages by using language through textual interactions (Maheswaran and Meyers-Levy, 1990). Particularly, to produce favorable responses, users need to use positive language exuding confidence, credibility, trustworthiness and optimism (Aktas et al, 2016; Avey et al., 2011; Davis et al, 2017; Martens et al, 2007). In line with this view, the previous research in branding area shows that when brands fail to communicate positive signals about their characteristics and quality, customers are likely to exhibit lower attitudes, quality perceptions and purchase intentions towards their products (Besharat, 2010). When applied in a P2P lending context, positive message framing to communicate information about the attributes of borrowers may be correlated to funding success by facilitating the evaluation of costly signals (like credit ratings) reflecting the qualities of borrowers (e.g., Anglin et al, 2018; Ahlers et al., 2015; Courtney et al, 2017).

2.2.1 Hypothesis Development

Trustworthiness is an individual attribute which reduces transaction costs and provide relational advantages in exchange relationships (Dyer and Chu, 2003). Trust can be either cognitive or affective. Cognitive trust emerges from the perceived ability and competence of the other party. Affective trust derives from positive feelings and caring motivations of the other party (Dowell et al., 2015). In this sense, positively framed message signaling trustworthiness can stimulate both cognitions and emotions in the trusting party (or trustor).

Öhman el at (2001) demonstrate that emotions communicated through language can drive attention by generating a number of positive sentiments towards an individual or a situation. Sentiment is a person's positive and/or negative emotional disposition towards another person or object. For example, sentiments of trustworthiness enable the trustor to establish emotional bonds with the trustee (Dowell et al., 2015). Thus, in general, positive sentiment may result in positive responses, whereas negative sentiment elicits negative reactions (Hsu et al., 2019). In the context of P2P lending, if the sentiment expressed in the description of the potential borrower and of the project conveys the notion that the borrower is trustworthy, then the project will receive attention from potential lenders which will translate into a higher number of received bids and possibly being funded. In particular, borrowers can enhance their trustworthiness by demonstrating positive and affective sentiments of concern and care in making timely repayments of loans to the lenders. Dorfleitner et al.'s (2016) study supports this view by observing that specific keywords (attached to certain sentiments) may have a positive influence on funding success of P2P loans. As a result, possible lenders will pay attention to those who use positively framed messages that emphasizes borrowers' trustworthiness in terms of their caring and concerning attitude for the investors.

Studies on message persuasion suggest that affective and cognitive states of persuasion are intertwined rather than separate (Homer and Yoon, 1992). On the one hand, affective framing of messages through positive language is observed to prompt the cognitive involvement of the receivers (Lee and van Dolen, 2015). On the other hand, messages triggering cognitive beliefs regarding trustworthiness can develop affective state of trust (Johnson and Grayson, 2005). In terms of cognitive effects, positively framed messages have a direct effect on cognitive beliefs regarding the trustworthiness of a party (Claeys and Cauberghe, 2014). Indeed, positive language has the potential to convince receivers about a sender's trustworthiness in terms of its ability and competence to achieve a particular goal (Newman et al, 2014), and control over outcome (Luthans et al, 2004). In line with this view, previous studies report that an individual's need for cognition or cognitive effort devotes more attention to rationally framed messages (Zhang and Buda, 1999). An emphasis on the ability and competence of a trustee would be interpreted as signals affecting cognitive beliefs and can therefore strongly impact the probability of funding success (Ahlers et al., 2015).

Framing language to communicate the message of trustworthiness is even more important in low-trust cultures such as China. At the personal level, "low trust" in the Chinese communities manifests itself in the lack of personal trust. Fukuyama (1995) attributes culture to the origin of "low trust" in the Chinese communities in which people can trust people they have a connection with. At the societal level, "low trust" also manifests itself in the lack of trust towards institutions and governing bodies. Ke and Zhang (2003) argue that underdeveloped institutions such as defense of private property rights and lack of market mechanisms underlie the prevailing low institutional trust in the Chinese communities. In such institutional environments, it is hard to trust people that are not known personally (Fukuyama, 1995; Putnam, 2000). In low-trust cultures such as China, investors need to be reassured about the trustworthiness of the borrowers, particularly in online

P2P settings, even more than in Western countries characterized by a high trust culture (Welter & Smallbone, 2006). Thus, positive message framing can play a greater role in triggering cognitive beliefs regarding trustworthiness and developing affective state of trust. Against this background, we posit that:

Hypothesis 1a (H1a). *Framing the message so that it suggests the borrower is trustworthy is positively related to the likelihood of a project being funded.*

Hypothesis 1b (H1b). *Framing the message so that it suggests the borrower is trustworthy is positively related to the number of bids a project receives.*

In an environment with pervasive asymmetric information, making decisions based on one set of signals – either costly or not - can be difficult. However, the presence of different types of signals may not solve the problem (Plummer et al., 2016) since taken together the two sets of signals may provide contradictory messages to receivers of the signals (Anglin et al, 2018). In such situation, receivers implicitly rank the quality and clarity of the signals they receive and decide which signal they decide to rely on. In general, receivers would tend to rely on costly signals more than on cheap signals due to their more reliable nature (Connelly et al., 2011). However, in the context of P2P lending, costly signals including individual credit scores (or ratings) can be complemented with the information on the quality of the potential borrower that cheap signals can provide. For example, Davis and Allison (2013) show that while costly signals are beneficial on funding success, cheap signals enhance these benefits by providing an effective form of communication between entrepreneurs and funders. Indeed, in noisy environments like P2P lending platforms, more information is required to make decisions, in which the complementary role of positive message framing emphasizing the trustworthiness of a borrower may enhance the effectiveness of costly signals such as credit rating in funding success (Anglin et al., 2018). In other words, positive message framing may strengthen the impact that other signals have on the likelihood of receiving funding (Anglin et al, 2018). We therefore suggest that:

Hypothesis 2a (H2a). *Framing the message so that it suggests the borrower is trustworthy positively moderates the association between credit rating of the borrower as a costly signal and the likelihood of a project being funded.*

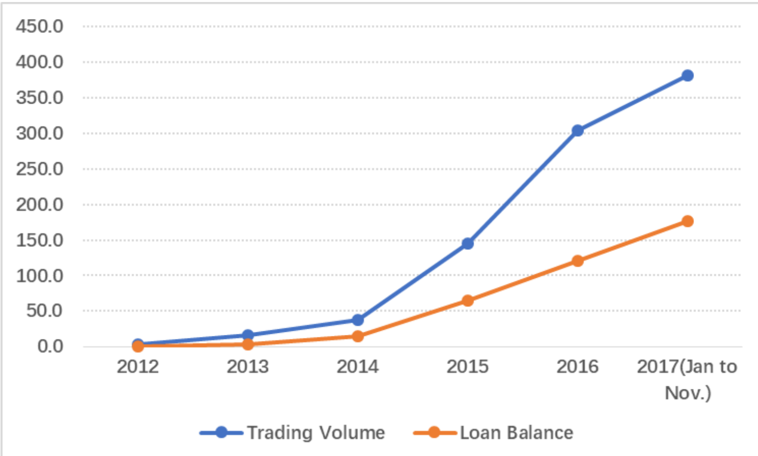
Hypothesis 2b (H2b). *Framing the message so that it the borrower is trustworthy positively moderates the association between credit rating of the borrower as a costly signal and the number of bids a project receives.*

2.3 Empirical analysis

2.3.1 P2P lending in China

Scholars define P2P lending as unsecured lending relationship that cuts off financial intermediary between investors and lenders by using online platform (Lin et al, 2013; Bachmann et al, 2011; Tao et al, 2017). In this industry, China is the world leader by any indicator. The first P2P platform, PPdai, was established in 2007 and the market began to explode in 2013. According to data collected by P2P consultant WDZJ.com, the trading volume of the Chinese P2P lending market expanded to 127 times its original size, from $3 billion in 2012 to $381 billion in November 2017. Meanwhile, total outstanding loans grew 221 times, from $0.8 billion in 2012 to $177 billion in November 2017 (Figure 1).

Figure 1 P2P Trading Volume and Loan Balance in China: 2012-2017



Three main factors explain the popularity of P2P lenders platforms in China (Huang, 2018; Mittal and Lloyd, 2016). The first one is the supply shortage in the formal financial markets, especially for small- and medium-sized enterprises (SMEs) and poor households. In China, banks overlook small borrowers (Fungacova and Weill, 2014). State-owned banks dominate the financial system, with a preference for lending to state-owned companies and the absence of a mature system for assessing consumer credit-risk adds to banks' reluctance to lend to individuals (Boyreau-Debray and Wei, 2005). Second, the regulatory environment has been very tolerant to alternative finance providers. Lack of regulation in the industry is explained by the belief that alternative providers could produce useful innovations while giving raise to limited risks due to their size (Wang et al., 2016). The third factor is the rapid development of internet-based startups, which are challenged to borrow from traditional sources of finance due to their limited resources and instable revenue stream (Salomon, 2018).

2.3.2 The Renrendai platform

For our empirical analysis we will use data extracted from an online P2P platform called Renrendai. This is one of the largest P2P platforms operating in China and several studies have already used their data (Mi and Zhu, 2017; Tao et al, 2017 and Yao et al, 2018). Renrendai was established in 2010 by three graduate entrepreneurs (Zhang, Li, and Yang) and it is the third largest P2P platform in China as of October 2018 according wzdj.com).

The loan application process set up by Renrendai is as follows. First, borrowers submit their application form with their national ID number and other personal information. They would need to specify the requested amount, the interest rate they would like to pay, the expected duration and purpose of the loan and any other information they find helpful to their application. The platform will verify the information submitted by applicants and assign a credit rating to the applicant varying between HR (high risk) and AA (very safe). The loan application process adopted by Renrendai is very similar among Chinese P2P platforms. First, like the majority of Chinese P2P platforms, Renrendai does not offer Web 2.0 functionality (Liu et al, 2018) implying that investors can only see static listing information and are not able to directly exchange information with the borrowers[2]. In addition to this, another key difference is that unlike developed economies that use a pure online lending process, Renrendai (and other Chinese P2P platforms) provide an offline authentication service to further mitigate asymmetric information that is originated by pure online authentication. In particular, they all use in-house credit ratings (generated using data submitted by the loan applicant at the moment of the application) to segment borrowers according to risk. At beginning of the operations, Renrendai adopted a pure online model similar to US platforms;

[2] This is different from what happens in other countries. For instance, investors in Prosper (a US based P2P platform) can communicate with the potential borrowers and potentially can elicit more information about their trustworthiness.

however, they changed their model after they merged with Ucredit[3] and started to trade as Youxin Financial. Renrendai explains their process of pre-loan risk control on their website as follows:

- Use of self-developed credit and risk analysis system;
- Verification manually conducted by themselves or their partner (Ucredit) of **every** loan application to ensure the authenticity of all information. In their website, they explain that they use the internet, telephone and other means they deem efficient to conduct what they claim "thorough and careful verification".

Credit rating is assigned by Renrendai. But they use two approaches to do it and bill the resulting applications as two different investment products. They offer borrowers the choice of whether to use their offline verification services or not. If the services are used, the application is billed as an offline verified investment product. Otherwise, it is billed as authenticity checked product. The respective processes are as follows:

- Option one: the application submitted online → documents desk checked → A credit rating assigned → the application up online for bidding
- Option two: the application submitted to a Ucredit's branch with a request for use of offline verification services → documents checked and verified offline → the verified application transferred to Renrendai → a credit rating assigned → the application up online for bidding.

All borrowers recommended by Ucredit will be assigned A class credit rating when their applications are listed online. Credit ratings of non-offline verified applications range more widely.

[3] Ucredit was founded in May 2011 in Shanghai by a team of entrepreneurs. Ucredit focuses on micro financing to individuals, and has a network of 300 branches in near 100 cities nationwide. Individuals can apply for micro loans to Ucredit in four ways: a) Online application through Ucredit website; b) Online application using WeChat APP; c) Application in branches and d) Application through Ucredit customer services hotline. Ucredit does not list any loan applications on their website for investors to bid, although they accept online applications. Ucredit currently focuses on two products: *Instant micro loans* that targets individuals who have a credit line of up to RMB 300,000 for the purpose of personal consumption, *Elite micro loans* that target civil servants, policemen, doctors, lawyers, and employees in large state-owned enterprises and banks. Micro loans to micro and small businesses do not appear to be their main focus anymore.

Renrendai charges borrowers initial service fees (one-off) 0%, 1%, 2%, 2.5%, 3%, 4% and 5% for AA, A, B, C, D, E, HR loans respectively. After that, they charge monthly management fees 0.55%, 0.60%, 0.65%, 0.70%, 0.75%, 0.80%, and 0.88% for AA, A, B, C, D, E, HR loans respectively. Interest rates vary in accordance with a borrower's credit rating, ranging from 6% to 24%. If the listing is unsuccessful, there is no fee for the applicants. Loans accessed through the Renrendai platform are all uncollateralized. The maximum duration of the loan is up to 3 years and the size of the loan ranges between 3000RMB and 500,000RMB. Renrendai guarantees the loan will be paid back to lenders by the end of the loan. To do so, it has a reserve fund to cover possible defaults and late payment. The fund is topped up constantly by the service fees charged. If the platform fails to collect back the loan, a collection agency will step in and the money eventually collected will be put it into the reserve fund.

The data for our empirical analysis is retrieved from the Renrendai platform and refer to all the listings between 1 Jan 2015 and 31 Dec 2015. We only focus on the loans that will be used to fund a start-up-related activity (based on the listing title)[4]. There are 43824 listings in our dataset. After deleting missing values, our dataset of made of 33028 listings. Among them, 9020 listings are fully funded. Each listing has the full set of information available to potential lenders. These include: (a) the loan terms such as interest rate, loan amount, loan duration, (b) credit rating assigned by the platform and (c) demographic information of the applicant such as gender, age, educational attainment, marital status, employment status and personal income range.

Why should some applicants prefer online applications only if they are more likely to be downgraded and less attractive to investors? Since Ucredit does not charge loan applicants for the use of their offline verification services, affordability of offline services is clearly not a factor.

[4] Listing titles vary substantially and include (among the others) medical expense, house purchasing, wedding ceremony preparation etc.

Three potential explanations can be accessibility, risk-taking, and self-confidence. First, Ucredit establishes their branches primarily in the first-tier cities (Beijing, Shanghai, Guangzhou and Shenzhen) and in some second-tier cities (Chongqing, Dalian, Chengdu, Ningbo). This means that they are less likely to conduct verification on applicants in cities they do not have branches. So, branch accessibility can be reasonably considered a factor that influences the take-up of offline services. Second, non-users of offline services are generally young (on average they are 30 years old) and may be more willing to take a chance with online services. Mature users may be more conservative than younger ones and thus are more likely to take up the offline services. So, risk-taking may be a factor. Third, as opposed to users of offline services, non-users are less likely to own a home property and a car and thus are less likely to have a mortgage and a car loan. As a result, the average loan they asked in the P2P platform is ten time smaller (RMB 5,279 against RMB 69,171). Indeed, their borrowing records suggests that non-users have very low success rate in loan applications anyway. For every ten applications they made, they succeeded at the rate of 0.17. In contrast, for users, for every 1.2 applications, the success rate was 1.007. Taken together, it is suggested that individuals who do not use offline services have no collateral and no credit history, implying that they cannot have access to loans. Hence, the offline services work like banks: individuals tend to avoid them if they know they have no chance of getting funds.

2.3.3 Empirical model and variables

In our model the propensity of a project to get funds (and the number of bids it attracts) is a function of a number of projects' characteristics (interest rate, amount requested and maturity) as well as the characteristics of the potential borrower namely its income, education and age. We add the indicator of trust and the number of words as our key regressors that allow us to test for the relevance of H1. The difference between the two models is in the dependent variable. Whether a loan is funded is dependent variable for Model 2 while the number of bids a listing obtains is the dependent variable for Model 3. Whether a loan is funded is a dummy variable, which takes one if the loan application is successful and zero otherwise. The second dependent variable is the number of bids which is proxied by the number of lenders who bid in this campaign. If H1 holds true, then the coefficients associated to the variable trustworthiness will be positive and significant. To test H2 we use similar specifications as above but we interact the indicator of trust with the applicant's credit rating (as assigned by the platform); if H2 holds true, the coefficient β_3 and γ_3 will be positive and significant.

Importantly, borrowers who would not choose offline authentication have different characteristics from those who would. Hence, it is highly likely that factors that influence whether borrowers use offline services could be correlated with the dependent variables of each equation creating a sample selection bias. To correct for such a bias, we use the Heckman selection model to estimate our model (Heckman, 1979). The Heckman model has two steps: the first step is to estimate the likelihood of borrowers choosing offline authentication by applying a probit model to whole sample including all listings with and without offline authentication. In the second step, the sample is restricted to the listings without offline authentication. Both dependent variables, offline authentication and funding success, are dummy variables and therefore we estimate a probit model with sample selection. STATA estimates the two equations simultaneously by using a maximum likelihood estimation approach (Andres, 2014).

The identification of the selection equation requires at least one variable that influences the choice of offline authentication but not affect investors' decision (Pham and Talavera, 2018). So, we create *Branch* which is equal to one if borrowers' location has an offline branch to do physical check and zero, otherwise. This is also a proxy for financial accessibility. People who live in some Chinese major cities are better off because in these cities, they easily get access to funding for their business. We conjecture that borrowers are more willing to do offline check if there is an offline branch because by doing so, the loan application is likely to be successful (Tao et al, 2017) and there is no additional cost. This variable is valid because investors' decision depends on other factors such as interest rate and duration. Other variables include the presence of a mortgage (1/0), car loan (1/0), age, income and degree. These variables are proxies for the presence of collateral which may influence the likelihood of having a project funded. Similarly, the model implies that borrowers with no mortgage and no car loan cannot have access to loans. Since the offline services work like banks, they tend to avoid them because they know they have no chance of getting funds. Our empirical model is as follows:

$$\text{Prob}(\text{NonOffline} = 1) = \alpha_0 + \alpha_1 \text{Mortgage} + \alpha_2 Carloan + \alpha_3 Branch + \alpha_4 Age + \alpha_5 Income + \alpha_6 Degree \qquad (1)$$

$$\text{Prob}(\text{FundingSuccess} = 1) = \beta_0 + \beta_1 \text{LnWordCount} + \beta_2 Trust + \beta_3 TrusCre + \beta_4 CreditRating + \beta_5 Interest + \beta_6 Duration + \beta_7 Ln\ Amount + \beta_8 Age + \beta_9 Income + \beta_{10} Degree \qquad (2)$$

$$\text{No of Bids} = \gamma_0 + \gamma_1 \text{LnWordCount} + \gamma_2 Trust + \gamma_3 TrusCre + \gamma_4 CreditRating + \gamma_5 Interest + \gamma_6 Duration + \gamma_7 Ln\ Amount + \gamma_8 Age + \gamma_9 Income + \gamma_{10} Degree \qquad (3)$$

(2) and (3) are the main outcome equations and (1) is the sample selection equation. We estimate our Heckman selection model using a Maximum Likelihood estimator.

Trust is our key independent variable in the outcome equations. We use the STATA function *ustrregexm* to search for words in the project description that are associated trust and trustworthiness. The Chinese words we select are “chengshi” (honesty), ”chengxin” (integrity), ”kekao” (trustworthy) , “xinyong” (credence) , “kaopu” (reliable) and ”xiangxin” (trust). If any of these words appears in the description of the project, the variable trustworthiness takes the value of 1; if it does not appear it takes the value of zero. Words of this kind are related to the drivers of ‘cognition-based trust’ as opposed to that of ‘affective-based trust’ (e.g. mutual understanding, personal relationship, etc.) (McAllister, 1995). McAllister (1995) found that cognition-based trust is a precursor of the development of affective-based trust. In addition, in an exploratory study examining the meaning of the language of trust as perceived by Chinese entrepreneurs, Tan and Chee (2005) found that the frequency count of factors of trustworthiness in words of ‘honest’, ‘integrity’, ‘sincerity’, ‘discreetness’ and ‘fairness’ facilitated the development of trust among Chinese entrepreneurs.

We choose the *CreditRating* by the platform as our moderator. Renrendai will refer to submitted documents and determine borrowers’ crediting rating from HR to AA. We code 0 for HR, 1 for E, 2 for D, 3 for C, 4 for B, 5 for A and 6 for AA.

Control variables include number of words used to describe the project, education, age, loan amount, loan duration and monthly income. A longer text is considered to be a signal of openness and transparency from the borrowers’ end of the transaction and therefore we expect that lenders are willing to fund loans that are well described and explained (Dorfleitner et al, 2016). In addition, the longer text, the more likely positive words could be included. Yet, if the description text is too wordy, the investors may not have enough patience to go through all the text. Since they only invest a small amount of money, it is not worth to cost much time (Dorfleitner et al, 2016). Taken together, we expect a positive association between the number of words and investors’ action.

Income indicates income range where n=0(less than 1000RMB), 1 (between 1000 and 2000RMB), 2(between 2001 and 5000 RMB),3 (between 5001 and 10000 RMB), 4(between 10001 and 20000 RMB), 5(between 20001 and 50000 RMB), and 6 (over 50000 RMB); *Degree* indicates the education level where n=0 (high school or lower), 1(junior college), 2(undergraduate), and 3 (master or above); *NonOffline* is coded one if the listing has not been verified through the offline authentication. Among the listings that were not funded, 22% of the project and applicants' description contain words that are related to trustworthiness while the proportion goes down to 3% for the projects that were funded. The length of the description of the projects varies among groups of projects: the average length of the non-funded projects is equal to 58 words while this goes up to 110 words among the funded projects. Interest rates for the two groups of projects tend to be different although they are in the region above 10%: more specifically it is equal to 12.6% for the projects that were not funded and 11.4% for the projects that were. Credit ratings for the projects that have not been funded are on average around HR while these go up to B for projects that have been funded. Virtually no funded project has been authenticated offline while around 71% of the funded projects have been authenticated offline. The average amount of the awarded loans varies between the two groups: it is equal to around 86500RMB for non-funded projects but it is equal to 59492RMB for the funded ones. The average proposed duration of the non-funded projects is 20 months while it goes up to 25 months for the funded projects. There is no difference in terms of the average income of the project applicants: it is between 5000 and 10000 RMB for both groups of projects. Table 1 describes the variables while the descriptive statistics are listed in Table 2.

Table 1. Variables Description and Expected Signs

Dependent Variables	Description	Expected Sign
Funded	Code one if the listing is successfully funded	Dependent Variable
Bids	Total number of investors in the campaign	Dependent Variable
Non Offline	Code one if the listing has not been verified through the offline authentication.	Dependent Variable
Independent Variables		
Trust	Code one if the borrower mentioned he/she is a trustworthy person in the description. Chinese words“chengshi”“chengxin““kekao”“xinyong“ “kaopu”“xiangxin“	Positive
Credit Rating	Renrendai has its own credit rating system from the highest risk (HR) to the AA rating. We code 0 for HR, 1 for E, 2 for D, 3 for C, 4 for B, 5 for A and 6 for AA.	Positive
Interest Rate	The interest rate on listing.	Positive/Negative
Months	The number of months the borrowers would like to pay back the loan.	Positive/Negative
Income	Variables indicating income range where n=0(less than 1000RMB), 1 (between 1000 and 2000RMB), 2(between 2001 and 5000 RMB),3 (between 5001 and 10000 RMB), 4(between 10001 and 20000 RMB), 5(between 20001 and	Positive

	50000 RMB), and 6 (over 50000 RMB).	
Age	The age of the borrower measured in years.	Positive
Degree	Variables indicating the education level where n=0 (high school or lower), 1(junior college), 2(undergraduate), and 3 (master or above).	Positive
Word Count	Number of words used to describe the project	Positive
Amount	Total Amount of Money that borrowers request	Positive/Negative
Mortgage	Code one if the borrower has the mortgage and zero otherwise.	Positive/Negative
Car loan	Code one if the borrower has a Car loan and zero otherwise.	Positive/Negative
Branch	Code one if there is an offline branch in the borrower's city and zero otherwise	Negative

Table 2. Summary Statistics

Not Funded Loans					
	Obs	Mean	Sd	min	Max
Word Count	24008	58.52	52.67	4	718
Credit Rating	24008	0.040	0.320	0	6
Trust (1/0)	24008	0.220	0.420	0	1
Amount (RMB)	24008	86500	110000	3000	500000
Age	24008	30.33	6.160	22	56
Degree	24008	0.830	0.780	0	3

Income	24008	3.060	1.170	0	6
Interest Rate(%)	24008	12.64	0.680	8	13
Months	24008	20.14	6.220	3	36
Non Offline(1/0)	24008	1	0.040	0	1
Bids	24008	0	0	0	0
Mortgage (1/0)	24008	0.118	0.322	0	1
Car loan (1/0)	24008	0.066	0.249	0	1
Branch (1/0)	24008	0.434	0.495	0	1
Funded Loans					
Word Count	9020	116.8	58.11	19	492
Credit Rating	9020	4.400	1.510	0	6
Trust (1/0)	9020	0.0400	0.190	0	1
Amount (RMB)	9020	59492	33715	3000	250000
Age	9020	37.16	8.430	20	62
Degree	9020	0.940	0.760	0	3
Income	9020	3.610	1.320	0	6
Interest Rate(%)	9020	11.44	0.960	8	13.20
Months	9020	25.14	10	3	48
Non Offline(1/0)	9020	0.290	0.450	0	1
Bids	9020	59.27	65.89	1	1165
Mortgage (1/0)	9020	0.354	0.478	0	1
Car loan (1/0)	9020	0.094	0.292	0	1

Branch (1/0)	9020	0.869	0.336	0	1

Note: *Trust* is coded 1 if the borrower mentioned he/she is a trustworthy/person in the description; RMB is Chinese currency and 1 USD=7 RMB in 2020.

As for the correlation among the variables, Table 3 reports the correlation coefficients among the variables. The correlation coefficients are generally significant while the value of the coefficients among the independent variables is not above 0.7 suggesting they can be used in the same model. We also check the multicollinearity among the independent variables using the VIF test (see table 3a) The average VIF is 2.24 and maximum VIF is 6.5. The results show no sign of multicollinearity given the rule of thumb is 10.

Table 3. Correlation coefficients

	Funded	WordCount	CreditRating	Trust	Amount	Age	Degree
Funded	1						
Word Count	0.432***	1					
Credit Rating	0.919***	0.465***	1				
Trust	-0.217***	0.050***	-0.248***	1			
Amount	-0.128***	0.027***	-0.082***	0.00700	1		
Age	0.406***	0.233***	0.432***	-0.118***	0.129***	1	
Degree	0.066***	0.084***	0.051***	0.021***	0.102***	0.00500	1
Income	0.198***	0.194***	0.210***	-0.049***	0.332***	0.253***	0.041***
Interest Rate	-0.574***	-0.345***	-0.579***	0.145***	0.146***	-0.225***	-0.035**
Months	0.287***	0.078***	0.368***	-0.096***	0.117***	0.188***	0.046***
Non Offline	-0.795***	-0.238***	-0.870***	0.224***	0.053***	-0.383***	-0.021**
Branch	0.390***	0.163***	0.421***	-0.085***	-0.035***	0.147***	0.035***
Bids	0.609***	0.264***	0.628***	-0.157***	-0.00400	0.296***	0.060***
Mortgage	0.272***	0.097***	0.289***	-0.059***	0.118***	0.190***	0.100***
Car loan	0.047***	0.00500	0.049***	-0.011**	0.077***	0.029***	-0.00800

Note: Lower-triangular cells report Pearson's correlation coefficients; * p<0.10, ** p<0.05, *** p<0.01

Table 3a.

Variable	VIF	1/VIF
CreditRating	6.5	0.153761
InterestRate	5.56	0.179714
Months	4.57	0.218799
LnWordCount	1.58	0.630975
LnAmount	1.46	0.686315
Income	1.37	0.728013
Age	1.31	0.761237
Branch	1.22	0.816368
Trust	1.18	0.843919
Mortgage	1.18	0.84968
Trust*CreditRating	1.06	0.941851
Degree	1.04	0.966134
Carloan	1.03	0.969032
Mean VIF	2.24	

2.4 Results

Table 4 shows the marginal effects of first step Heckman selection model. The dependent variable, *Non Offline*, is a binary variable indicating whether a borrower engaged in offline authentication. *Non Offline* takes one if the borrowers didn't choose offline authentication, because in second step, we only focus the borrowers without offline authentication. Model 1 is the baseline model: the independent variables include an offline branch dummy, *Branch*, and measures of borrowers' different loan variables such as *Mortgage* and *Carloan*. Model 2 introduces borrower's *Age*, *Income* and *Degree* obtained as additional independent variables choice.

Table 4. Selection Equation

DV: NonOffline	(1)	(2)
Mortgage	-0.184***	-0.139***
	(-51.95)	(-40.19)
Carloan	-0.046***	-0.048***
	(-7.28)	(-8.49)
Branch	-0.504***	-0.453***
	(-39.03)	(-40.19)
Age		-0.011***
		(-63.84)
Degree		0.001
		(0.52)

Income		0.002**
		(2.05)
N	33028	33028
Pseudo R-square	0.327	0.421
Log Likelihood (LL)	-10945.278	-9426.570
Deviance (-2LL or Chi square change vs. Model 1)		-3037.42***

The Table shows the average marginal effect of the variables on the NonOffline. Z statistics in parentheses;
* p<0.10, ** p<0.05, *** p<0.01

In Model 1, if there is an offline branch in borrowers' location, applicants are more willing to conduct offline authentication, so the coefficient of *Branch* has a negative sign on *Non Offline*. Loan history variables, mortgage and car loan show a similar behaviour so there is a negative relationship between loan history variables and *Non Offline*. In Model 2, *Income* has a positive and significant impact on *Non Offline*. There is a negative association between *Age* and *Non Offline*, which means young borrowers are less likely to opt for offline verification. The value of the Chi square (-3037.42, p<0.01) at the bottom of table 2 implies that Model 2 has a better fit. Hence, we use Model 2 as selection model.

The results of second step are shown in Table 5. Generally speaking, the Heckman model is well-suited because the P-Value of the Wald test from *Heckprobit* is less than 0.05 for all the model specifications. We use hierarchical multiple regression analysis for hypothesis testing. The Model 1 consists only three key independent variables. Model 2 and Model 3 add demographic and listing characteristics respectively. Lastly, we add an interaction term, *Trust*Credit Rating*, which is the interaction between trustworthiness and credit rating, in Model 4 to test the moderating effect of trustworthiness.

The coefficients associated to the variable Trustworthiness are positive and significant across the different specifications and in general it supports H1a. In other words, applicants who use words that are associated to trust tend to be fully funded. When the variable Trustworthiness is interacted with the Credit Rating variable (Model 4), the interaction term, *Trust*CreditRating*, is still positive and significant i.e. for applicants that use words associated to Trust, there exists a net positive association between Credit rating and funding success (confirming H2a). Norton et al (2004)' command, *inteff*, cannot be applied to this study as we use *Heckprobit* to estimate the models. However, we have calculated the marginal effects of the interaction term manually using STATA. As a result, the marginal effects are shown in all Tables.

Results suggest that borrowers who do not choose offline authentication, all things being equal, are more likely to be funded if they frame their message around the concept of trust. Consistent with the findings of Lin et al. (2013), the probability of funding success increases with the credit rating ($P<0.01$) in all model specifications. Natural logarithm of the number of words (*LnWordCount*) is also positive and statistically significant in all model specification.

Demographic variables, *Age*, *Degree* and *Income* are in general positive and significant (Tao et al., 2017). As for the characteristics of the applicants, the older the applicant, the higher its income and the more qualified it is more likely it is to obtain funds. In terms of the other control variables, the amount of the requested loan has a negative impact on funding success. The interest rate is positively associated to the funding success. Variables associated to the characteristics of the listings (such as *InterestRate, Months and LnAmount),* are highly significant and the marginal effects are in line with our expectation. Higher interest rates attract lenders and increase probability of funding success. *LnAmount* is negative and significant, implying that P2P investors tend to bid on small loans and diversify the risk. When the loan amount increases, lenders are suspicious about the borrowers' ability to repay the loan, thereby decreasing the probability of funding success (Herzenstein et al, 2008). In contrast with the finding of Tao et al (2017), we find entrepreneurial loan duration has a significantly negative impact on funding success, suggesting lenders perceive long-term entrepreneurial loans are risky. Lastly in Model 4, we test the moderating effect of the credit rating by interacting trust language with credit rating. This interaction term, *Trust*CreditRating*, is positive and highly significant (P<0.01), confirming our hypothesis.

Table 5. Outcome Equation: Probability of being funded.

DV: Funded	(1)	(2)	(3)	(4)
LnWordCount	0.003**	0.002*	0.010***	0.011***
	(2.37)	(1.78)	(5.57)	(6.00)
CreditRating	0.056***	0.056***	0.062***	0.061***
	(41.12)	(41.25)	(28.55)	(28.04)
Trust	0.007***	0.006***	0.006**	-0.004
	(3.12)	(2.98)	(2.25)	(-1.23)
Age		0.000***	0.001***	0.002***
		(3.92)	(9.42)	(9.28)
Degree		0.008***	0.016***	0.016***
		(7.02)	(10.09)	(9.83)
Income		0.001	0.018***	0.018***

		(0.72)	(13.56)	(13.43)
LnAmount			-0.052***	-0.054***
			(-21.67)	(-21.36)
InterestRate			0.017***	0.014***
			(5.25)	(4.29)
Months			-0.003***	-0.002***
			(-6.98)	(-6.10)
Trust*CreditRating				0.036***
				(7.30)
N	33022	33022	33022	33022
P-Value Wald test of Indep.	0.00	0.03	0.04	0.04

The Table shows the average marginal effect of the variables on the funding probability. Z statistics in parentheses; * $p<0.10$, ** $p<0.05$, *** $p<0.01$

In the last batch of results, we test whether message framing matters for the number of bids associated to each project. Empirically, the model is similar to the one estimated above although the dependent variable is now a continuous one (i.e. not a dummy variable). The results are presented in Table 6. While the signs and the coefficients of the control variables are similar to those presented in Table 5, our key independent variable is not significant jointly with the interaction term suggesting that H1b and H2b are not verified.

Table 6. Outcome Equation: Number of Bids

DV: Bids	(1)	(2)	(3)	(4)
LnWordCount	0.460***	0.364***	0.398***	0.387***
	(3.30)	(2.59)	(2.84)	(2.76)
CreditRating	9.278***	9.203***	7.226***	7.256***
	(106.53)	(103.74)	(48.30)	(47.91)
Trust	-0.197	-0.190	-0.143	-0.082
	(-0.91)	(-0.88)	(-0.66)	(-0.37)
Age		0.018	0.025*	0.026*
		(1.30)	(1.78)	(1.81)
Degree		0.445***	0.447***	0.453***
		(4.05)	(4.06)	(4.11)
Income		0.274***	0.342***	0.343***
		(3.65)	(4.15)	(4.16)
LnAmount			-0.228**	-0.230**
			(-2.20)	(-2.22)
InterestRate			-5.025***	-4.989***
			(-15.99)	(-15.81)
Months			0.464***	0.459***

			(13.10)	(12.91)
Trust*CreditRating				-0.676
				(-1.25)
N	33022	33022	33022	33022
P-Value Wald test of Indep.	0.00	0.06	0.06	0.06

The Table shows the average marginal effect of the variables on the Number of Bids. Z statistics in parentheses; * p<0.10, ** p<0.05, *** p<0.01

2.5 Robustness Check

2.5.1 Sentiment Analysis

Sentiment analysis is able to identify and mine positive/negative opinions and emotions that are derived from the text (Wilson et al., 2005). The words in borrowers' loan description contain both positive and negative sentiment. Since the objective of this study is to mine the way that borrowers ask for a loan by analyzing their positive loan descriptions, it is appropriate to identify borrowers' sentiment orientation and extract the positive description (Zuo et al., 2019). Following Zuo et al (2019), Guo et al (2019) and Zhang et al (2017), we conduct sentiment analysis by using SnowNLP, a Chinese natural language processing library in Python. This library has similar functions as TextBlob (a Python library for English textual data), such as word segmentation, part-of-speech tagging, text abstraction, and sentiment analysis. For each input loan description, SnowNLP can predict the sentiment orientation of the text by built-in Chinese dictionary. The dictionary is constructed by SnowNLP developers[5]. The developers collect sentences from microblogs and online reviews and define these sentences as positive or negative languages. The precision rate of SnowNPL to identify emotional tendencies of social media texts is over 80% (Zhang et al, 2018). The accuracy of TextBlob in Tweets is 76% (Hasan et al., 2018)

Although SnowNLP allows users to apply their own dictionaries, to eliminate self-selection bias derived from using self-chosen dictionary, we choose to use built-in dictionary. In addition, as borrowers post their words online and online words-of-mouth shares similar unique features[6], SnowNLP built-in dictionary is well-suited for our analysis. In fact, the algorithm that SnowNLP uses is a Naive Bayes algorithm, a probability model for binary classification. Hence, SnowNLP output value is between 0 and 1 which implies the probability of the description being positive emotion. The higher value indicates the higher probability of positive sentiment. For example, one description is as follows ‘‘I found an investment opportunity so I need funding. My monthly income is RMB 3000. I have a house and good borrowing history. I will use my salary and the return from the investment to pay back the loan.’’ The emotional value evaluated by SnowNLP is 0.999, which means the probability of being positive emotion is 99.9%.

We use two loan description to show the degree to which SnowNLP can predict sentiment index of a negative prefix such as not before positive words. For example, the sentiment index of “I have some spare time after work so I would like to apply my first loan here to do some small business. I have no idea whether it is trustworthy or not trustworthy (in English, we may omit last trustworthy but in Chinese, we keep it) so I just try a small amount of money. I am

5 For an open source Python package in sentiment analysis, see https://github.com/isnowfy/snownlp for details. The positive/negative language dictionaries and detailed codes to calculate sentiment index are public on the developers’ website.

6 Yang (2007) suggests internet community enables language innovation so their writing system is different from other places. For example, huichang (灰常) is for feichang (非常), which means “very much”.

working at Changsha and have bought a 150 square feet flat in my hometown." is 0.37 while the probability of "I am a civil servant, working at local Justice Bureau and I am deputy head of the department. I have a stable job with monthly income of RMB 8000. I don't have bad habits. The purpose of this loan is to invest in a shop but there is no enough money. I am a trustworthy person and a veteran cadre at my workplace. I can make sure to repay the loan on time" is 0.97. From these examples we can see that SnowNLP is able to identify a negative prefix and treat them differently.

Table 7 shows the results of sentiment analyses. *Sentiment* is a continuous variable, predicted by SnowNLP directly. Model (1) shows it is positive and significant at 10% level. *Sentiment*CreditRating* is an interaction between *Sentiment* and *CreditRating*. Model (2) shows *Sentiment*CreditRating* is positive and significant at 1% level. To further confirm our results, following Zhang et al (2017)'s procedure, we first set the threshold to 0.6. If output value for loan description is higher than 0.6, we classify them as positive description. When the sentiment score is lower than 0.6, Zhang et al (2017) consider it neutral or negative. So, *Sentiment0.6* is a dummy variable, which takes one if loan description shows a positive sentiment and 0 otherwise. Model (3) and (4) show the results of *Sentiment0.6* and its interaction term with *CreditRating* respectively. Both are positive and significant at 5% level. In an additional robustness test, we set our threshold to 0.7, which means only if the positive possibility is higher than 70%, we consider it positive sentiment. This criterion is more restrictive than the previous one. The results are shown in Model (5) and Model (6) and are qualitatively the same.

Table 7. Sentiment Analysis

DV: Funded	(1)	(2)	(3)	(4)	(5)	(6)

LnWordCount	0.011***	0.013***	0.011***	0.013***	0.011***	0.012***
	(6.22)	(5.87)	(6.22)	(5.90)	(6.18)	(5.87)
CreditRating	0.063***	0.068***	0.063***	0.066***	0.062***	0.066***
	(27.14)	(27.15)	(27.60)	(25.62)	(27.70)	(25.10)
LnAmount	-0.053***	-0.065***	-0.053***	-0.061***	-0.052***	-0.059***
	(-20.55)	(-14.95)	(-20.84)	(-14.98)	(-20.94)	(-15.04)
Age	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(9.31)	(8.60)	(9.33)	(8.59)	(9.33)	(8.58)
Degree	0.016***	0.021***	0.016***	0.019***	0.016***	0.019***
	(10.00)	(8.99)	(10.22)	(8.99)	(10.04)	(8.99)
Income	0.018***	0.023***	0.018***	0.021***	0.018***	0.021***
	(13.23)	(11.17)	(13.33)	(11.18)	(13.35)	(11.18)
InterestRate	0.018***	0.022***	0.018***	0.021***	0.018***	0.020***
	(5.31)	(5.29)	(5.32)	(5.25)	(5.32)	(5.25)
Months	-0.003***	-0.004***	-0.003***	-0.003***	-0.003***	-0.003***
	(-7.00)	(-6.82)	(-7.02)	(-6.74)	(-7.02)	(-6.70)
Sentiment	0.006**	0.002				
	(2.02)	(0.60)				

Sentiment*CreditRating		0.013***				
		(3.22)				
Sentiment0.6			0.005**	0.003		
			(2.27)	(1.16)		
Sentiment0.6*CreditRating				0.008***		
				(2.59)		
Sentiment0.7					0.004*	0.002
					(1.89)	(0.88)
Sentiment0.7*CreditRating						0.007**
						(2.33)
N	33028	33028	33028	33028	33028	33028
P-Value Wald test of Indep.	0.04	0.04	0.04	0.04	0.04	0.04

The Table shows the average marginal effect of the variables on the funding probability. Z statistics in parentheses; * $p<0.10$, ** $p<0.05$, *** $p<0.01$

2.5.2 Subsample Analysis

In addition, we conduct subsample analysis to check the robustness of our findings. We divide our sample into two subsamples based on borrowers' location, i.e. the coastal region and the rest of China. The coastal region, including Heibei, Beijing, Tianjin, Shandong, Jiangsu, Zhejiang, Shanghai, Guangdong, Hainan and Fujian, is China's growth hub as opposed to the rest of the country. The results are shown in Table 8. The Model (1) and (2) are the coastal-provinces group. The coefficient of trust is not significant in this group while the interaction term of trust and credit rating is positive and significant. Non-coastal-provinces group is shown in the Model (3) and (4). Both trust and the interaction are positive and significant in this group. It is noteworthy that we cannot reject the test of independent equations for all the subsample analysis. This means the selection equation and outcome equation can be estimated separately and our finding still holds.

Table 8. Subsample Analysis

DV: Funded	(1)	(2)	(3)	(4)
LnWordCount	0.012***	0.013***	0.007***	0.009***
	(4.40)	(4.59)	(3.24)	(3.68)
CreditRating	0.063***	0.062***	0.059***	0.059***
	(21.40)	(20.93)	(18.44)	(18.27)
Trust	0.002	-0.007	0.010***	-0.001
	(0.51)	(-1.57)	(2.65)	(-0.29)
LnAmount	-0.056***	-0.058***	-0.048***	-0.052***
	(-16.38)	(-16.21)	(-14.09)	(-13.86)
Age	0.002***	0.002***	0.002***	0.002***
	(5.96)	(5.87)	(7.69)	(7.64)
Degree	0.013***	0.012***	0.020***	0.021***
	(5.64)	(5.34)	(8.45)	(8.41)
Income	0.021***	0.022***	0.013***	0.013***

	(11.11)	(11.02)	(7.05)	(6.98)
InterestRate	0.012***	0.009**	0.020***	0.018***
	(2.71)	(1.96)	(4.26)	(3.72)
Months	-0.003***	-0.002***	-0.003***	-0.002***
	(-4.77)	(-4.07)	(-4.80)	(-4.32)
Trust*CreditRating		0.035***		0.038***
		(4.83)		(5.60)
N	16234	16234	16788	16788
P-Value Wald test of Indep.	0.63	0.70	0.94	0.91

The Table shows the average marginal effect of the variables on the funding probability. Z statistics in parentheses; * p<0.10, ** p<0.05, *** p<0.01.

2.6 Discussion

2.6.1 Theoretical Implications

The findings of this study demonstrate that a positively framed message has a significant influence on funding success but not on the number of bids. Our results try to reconcile contrasting findings in the context of online reviews (e.g. Ludwig et al., 2013; Salehan and Kim, 2016) by showing that message framing matters in the case of Internet-mediated transactions. Specifically, this study shows that the content of a borrower's message and how it is framed using language is positively correlated to funding decisions but not to the number of bids. Theoretically, we have combined message framing theory with signalling theory to show that language can be a powerful signalling device even if the message is mediated by technological devices (like platforms or online sites).

Our findings are also consistent with research in marketing which has shown that framing has a much stronger effect when consumers have little or no related product experience (Chang, 2007). In this sense, our findings imply that message framing can be useful when investors have no familiarity with potential borrowers which are working towards obtaining funds on P2P lending platforms. In addition, our study contributes to the studies showing the ineffectiveness of cheap signals in generating new funding streams (e.g. Chen et al., 2009). In P2P lending context, investors may hold more limited knowledge and skills for assessing the various costly signals associated with qualities of a borrower, which may justify the importance of language as a cheap signal in P2P lending decisions. In this way, our study also builds on the main tenet of traditional signaling theory which suggests that cheap signals make it impossible to create a separating equilibrium between better and worse signalers (Balvers et al., 2014).

Importantly, our findings show the complementary role of costly and cheap signals in generating funding success. In particular, they reveal that costly signals including individual credit scores (or ratings) are complemented with cheap signals (such as positively framed messages) to generate insights into the quality of the potential borrowers. This finding contributes to the literature examining external signals in isolation, and thus has responded to the call for research on the interaction of different signals on certain performance outcomes (Anglin et al., 2018). While some studies suggest that observable actions (i.e. costly signals) are mostly considered to be a more credible signalling mechanism compared to pure words (Cheung et al., 2014). Our finding is consistent with the limited number of studies which suggest that while costly signals are beneficial on funding success, cheap signals enhance these benefits by providing an effective form of communication between entrepreneurs and funders (Davis and Allison, 2013). Such communication is particularly important in virtual environments, in which impression formation based on face-to-face interactions and experiences is lacking. The costly signals reduce the risks of cheating and any costs associated with misleading signals (Connely et al., 2011). However, the costs of relying exclusively on costly signals and assessing them may be high. As such, the complementary use of cheap signals may be cost-effective for senders of such signals.

2.6.2 Implications for Practitioners

This study suggests that message framing as a cheap signal can be effective in enhancing the likelihood of a project being funded. In particular, messages framed in a way to signal trustworthiness of potential borrowers have greater likelihood for funding success. Potential borrowers signaling messages for P2P lenders need to frame their messages with words such as honesty, integrity, credence, and reliable which signal their trustworthiness quality. Besides trustworthiness, potential borrowers can also display other qualities through language to attract investors and increase the potential for funding success. For example, signaling words associated with the agility, proactiveness, ambiguousness, empathy and network size of a potential borrower may support lenders' decisions on whether the borrower is worth investing. In this sense, message framing can be very relevant.

Furthermore, as suggested by our findings, to enhance the odds of funding success, costly signals such as information on the credit rating of a potential borrower can be complemented by suitably framed message embedded in the description of the project. This is because, when a borrower communicates their credit rating through objectively assessed evidence, their continuous emphasis on their credit history may only have a supplementary role, and thus have marginal additional influence on their funding success. If borrowers seeking for funds can enhance their communication mode through the use of positive messages, their qualities, which can be objectively assessed, would be supported through language, and therefore increase the likelihood of being funded.

2.6.3 Limitations of the Study and Future Research

This research has only focused on the use of language as a cheap signal on funding success. Future studies can focus on how message framing reflecting the qualities of potential borrowers can be complemented with other types of cheap signals such as message length and the format of writing to influence the funding success (Li and Zhan, 2011). In addition, this study only examined the role of message framing as a cheap signal that can motivate investors to fund the project. Future studies can investigate whether negative language may attract the attention of investors and how it may affect their decisions. For example, studies can investigate whether negatively framed messages explaining the risks of funding other types of borrowers competing for the same investor's resource or the negative messages about the competitors may be effective in attracting the attention of investors and enhance the odds of funding success. Specifically, future studies can use sentiment analysis to understand how creating negative emotions about the competing actors may affect a borrower's success of obtaining funds. Importantly, this study focused on borrower-related issues at the individual level from only the borrower perspective. Future research can use a multi-level study to analyze how individual-related and environmental issues, from both borrower and investor perspective, may affect funding performance of potential borrowers. Finally, this study only focused on the extent to which message framing may moderate the effect of credit rating of a borrower as a costly signal on their funding performance. Future studies can investigate the moderating effects of other types of cheap signals such as entrepreneurship orientation of a borrower on various types of costly signals including their education, credit history and existing resources. An important issue to highlight here is the extent to which these results can be generalized to other platforms and more importantly, whether it is possible to conduct similar studies that use data from different platforms. Crucially, P2P platforms in China use different business models and in this respect merging data from several platforms cannot offer useful insights as in reality the data generating processes are different because of the way the platforms are run. In this respect, a useful exercise is not so much to increase the volume of data analyzed by merging different datasets but to ascertain whether in platforms with different business models, message framing is as important as in the Renrendai platform.

2.7 Conclusions

This study has examined the role of message framing suggesting that the borrower is trustworthy in their funding success in terms of the likelihood of a project being funded. In addition, the study has investigated the extent to which it moderates the effect of credit rating information of a borrower as a costly signal on their funding success. The study has confirmed that positive message framing emphasizing the trustworthiness of a borrower is associated to funding success. The study also found that message framing and costly signals (like credit ratings) complement each other.

Chapter 3 Psychological distancing and Language intensity in P2P lending

Abstract

The study bridges the P2P lending literature and psycholinguistics literature and set out to explain how psychological distancing manifested by linguistic styles can influence lenders' decision on P2P funding campaign. We find that linguistic styles related to psychological distancing have a negative impact on P2P funding success. Moreover, the language intensity tends to strengthen the negative relationship between psychological distancing and funding success. This finding is consistent with psycholinguistics literature which suggests that psychological distancing is associated with negative interpersonal outcome (Simmons et al, 2005; Revenstorf et al, 1984). Specifically, the number of "you" and the number of negations used in borrowers' description are negatively related to the willingness of the lender to support the funding campaign. The intensive language negatively strengths the relationship between the funding performance and number of "you" but does not apply to number of negations.

3.1 Introduction

P2P lending has become an important source of funds for consumers and entrepreneurs. Consumers can use the funds for education, house purchasing or wedding ceremony. It is also important for entrepreneurs to fund their startup. This is particularly important in China because the banks are more willing to provide loans to State Owned enterprise and private business finds difficult to borrow money from banks (Fungacova and Weill, 2014).

This study examines the effect of psychological distancing arised from communication on the funding campaign. More specifically, we investigate whether psychological distancing in communication discourages P2P lending contributions. We selected this setting for several reasons. Firstly, P2P lending is the most important form of crowdfunding (Saiedi et al., 2020). Second, a successful funding campaign requires support from the crowd. Psychological distance in borrowers' communication affects the perceptions of the crowd and therefore, influences the campaign outcome. A sense of closeness by the crowd may promote the contributions. There are four dimensions of psychological distance including spatial distance, social distance, temporal distance and hypothetical distance, which suggests mind goes away from what is in front of right now. They are different dimensions of psychological distance because these constructs of consideration makes people feel psychologically distant (Maglio 2020). If something is psychologically distant (e.g., far away in temporal, spatial , probabilistic or social distance), people feel psychological distancing in a subjective sense (Maglio 2020).

Given the fact that P2P lending takes place in online settings, communication between lenders and borrowers is a challenge. Unlike offline interactions, narratives of the borrower plays a key role in online communication process. Anglin et al (2018) suggest the backers who support the funding campaign form assessments or make judgments relying on the narrative. The entrepreneurship literature shows how the entrepreneurs form the message matters as the narrations help them to acquire essential resources by setting expected outcomes and conveying value (Huang et al., 2020; Martens et al., 2007; Chen et al., 2009). Therefore, researchers have showed a great deal of interests in the role of language in crowdfunding and entrepreneurial finance literature (Defazio et al., 2020; Huang et al., 2020; Parhankangas and Renko, 2017; Larrimore et al., 2011).

A stream of literature has investigated the association between language and funding outcome in different types of crowdfundings including reward-based crowdfunding (Parhankangas and Renko, 2017; Anglin et al., 2018), equity-based crowdfunidng (Block et al., 2018) and P2P lending (Herzenstein et al., 2011; Allison et al., 2013; Majumdar and Bose, 2018). In different funding campaigns, borrowers' narratives normally focus on various aspects. In reward-based crowdfunding, the narratives often center on new product development plan[7], while in P2P lending, borrowers tend to stress on evidence-based elements because personal details and loan characteristics can appeal more lenders (Lee et al., 2019; Allison et al., 2013). Nevertheless, most of extant studies mentioned above focus on the content of the language. However, a few studies related to linguistic styles, measured by use of some style words found lenders form perceptions about prospective borrowers based on the style words and therefore style words can actually influence lenders' decision (Herzenstein et al., 2011; Larrimore et al., 2011).

Although some aspects of linguistic styles have been studied, the linguistic styles that are related to psychological distancing are scarce with one exception of Parhankangas and Renko (2017). We follow Parhankangas and Renko (2017) and define Psychological distancing is to the extent to which people distance or remove themselves away from the topic being discussed. They discuss psychological distancing and crowdfunding success for social entrepreneurs. The construal-level theory (CLT) provides a theoretical framework to understand psychological distancing (García et al., 2020). The CLT suggests the consequences of psychological distancing is reduced affective concern and people have a weaker affective response to distant stimuli (Williams et al., 2014). To build personal rapport from the crowd, it is advised that entrepreneurs' use of language should avoid psychological distancing (Parhankangas and Renko, 2017). Due to different focus of entrepreneurial narratives in P2P lending, it is necessary to investigate how P2P lenders react to perception of psychological distancing. Moreover, the language intensity that related to tone of communication can also affect the funding outcome (Han et al., 2018). Language intensity has been first studies since the seminal work of Bowers (1963). Following Clementson et al.(2016) and Bowers (1963), we define

[7] The plan usually tells the investors within three or six months, they can deliver their products to the market.

language intensity[8] as language implying direction and degree of distance from neutrality[9]. Drawing on language expectancy theory (LET), we suggest language intensity moderates the effect of psychological distancing on funding success. To date, researchers have studied these two factors in isolation. Yet, linguistic styles and language intensity are actually inseparable. How tone of communication complements linguistic styles to form the perception is still unknown even though lenders receive both together. In addition, scholars tend to pay attention to how positive linguistic styles improve likelihood of funding success. Extant literature largely ignores the negative effect of some linguistic styles such as psychological distancing on funding outcomes (Huang et al., 2020). Given the fact that the important role of psychological linguistic styles on online context (Ludwig et al., 2013; Peng et al., 2004) and there are still rich features that can be studied (Lee et al., 2019), we try to explore how psychological distancing affects P2P funding outcomes and how language intensity strengthen this relationship.

The CLT provides a theoretical framework to understand psychological distancing (Garcí a et al., 2020). To examine the effect of psychological distancing on P2P lending, we draw upon construal-level theory (CLT) as a theoretical base (Trope and Liberman 2003, 2010; Williams et al., 2014). In addition, Language expectancy theory (LET) posits language intensity interacting with other factors will determine persuasion power of the message (Burgoon et al., 2002). Hence, drawing on LET, we test the moderating role of language intensity on the relationship between psychological distancing and P2P funding performance. We use data from one of the leading Chinese P2P platform, Renrendai, to perform our empirical analysis. To alleviate sample selection bias, we also adopt a Heckman two stage model. The results show

[8] High intensity language includes emotion-laden words such as horrible and excellent, which suggests the extreme position and deviation from the neutrality (Bowers, 1963).
[9] There are two linguistic ways to deviate from the neutrality: 1) directness toward the audience and 2) use of emotion-laden words (Bradac et al., 1979)

that the psychological distancing in the description of the project negatively predicts the funding success. The result is consistent with psycholinguistics literature that suggests psychological distancing is associated with negative interpersonal relationship. We also contribute to the literature by extending CLT to an emerging lending market. Moreover, language intensity partial enhances the negative effect of psychological distancing on funding success. We contribute to the literature by finding the empirical support for LET.

3.2 Theoretical background and Hypothesis Development

CLT proposes an individual's perception of psychological distance to a stimulus (a person, an event or an object) determines the stimulus is mentally represented or construed. If the psychological distance is high to a stimulus, individuals tend to form an abstract of this stimulus that focuses on its general meaning (Trope and Liberman 2003, 2010). In addition, the CLT also argues that the consequences of psychological distancing is reduced affective concern (Williams et al., 2014). Transactions online in particular are more likely to be driven by affective concerns (Han et al., 2007). Maglio (2020) posits the mind and the body join together in dealing with affective information while psychological distance has an increasingly nuanced relationship with the affect. A sense of psychological closeness promote the affective response of the transaction, while, psychological distant mindset undermines affective elements (Williams et al., 2014). For example, when making judgments or choices, such as feelings of satisfaction with a specific product or feelings of empathy that promotes charitable giving, psychological distance can weaken the affective reactions to these events (Williams et al., 2014). Moreover, Homer and Yoon (1992) suggest that affective and cognitive states of persuasion are complementary rather than separate. This implies that affective elements may influence cognitive response which emerges from the perceived ability and competence of the other party (Huang et al., 2020; Lee and Van Dolen, 2015). In sum, based on CLT, we argue psychological distance can discourage the P2P lending transactions between the parties.

The words we use on a daily basis consist of both content words and style words (Abe, 2011; Toma and D'Angelo, 2015). Content words such as adjectives, verbs and nouns can carry much of the meaning, while style words or linguistic style[10] are related to how information is conveyed. Linguistic style are made of articles, auxiliary verbs, prepositions, conjunctions, pronouns and negations (Tausczik and Pennebaker, 2010; Toma and D'Angelo, 2015). Style words account for 55% of total words we normally use although they only represent 0.04% of all words (Pennebaker, 2011; Tausczik and Pennebaker, 2010). Moreover, human's brain processes content words and style words differently (Miller, 1995).

Linguistic style is more likely to be associated with the measures of individuals' social and psychological worlds than content (Pennebaker and Chung, 2013; Tausczik and Pennebaker, 2010), as style of communication may contribute to appeal to an audience for communicators (Parhankangas and Renko, 2017). Extant research shows that the frequency of certain style words is associated with how we are perceived by others and has an impact on outcomes such as academic performance and funding campaign of social entrepreneurs (Robinson et al.,2013; Parhankangas and Renko, 2017; Fausey and Boroditsky, 2010). Furthermore, Larrimore et al (2011) find stylistic features of messages can be effect in assessment of source trustworthiness especially in an online environment where social cues such nod and smile about the information source are less reachable. In an analysis of linguistic style of Chinese social media, intensive use of negations[11] can attract great number of readership (Wang, 2019). We therefore expect that P2P lenders are also sensitive to the use of linguistic styles of entrepreneurs.

[10] Style words or linguistic style are interchangeable in this study.

[11] For example, "No matter you believe or not, I do not believe." There are three negations in this sentence.

It is possible that entrepreneurs disconnect/connect to their potential lenders on an emotional level by unconscious use of certain style words that influence psychological distance because certain language use can distance away relationships (Tausczik and Pennebaker, 2010). In addition, research suggests that psychological distance is embedded in language use (Nook et al., 2017). The use of style words implies temporal and social "present" such as present-tense verbs and first person pronouns (Mehl et al., 2013; Pennebaker and King, 1999). Toma and Hancock (2012) suggest linguistically, psychological distancing manifest itself by a decrease in first person pronouns and an increase in negations. Psychological distancing may alienate the entrepreneurs who launch the funding campaign on a P2P platform from their crowd (Parhankangas and Renko, 2017). Similarly, Mehl et al. (2013) suggest considering psychological implication of language use in different communication context is important. For example, Tidwell and Walther (2002) find that interactive strategies that commonly employed to close the psychological distance in face-to-face context are less effective in online settings. So, people are likely to pay attention to the subtle but important information that is conveyed by linguistic style available online because of lack of common cues to get to know each other (Flanagin, 2007).We argue that avoiding psychological distancing is more essential for entrepreneurs in P2P platform. Entrepreneurs' ability to payback loan relies on the cash inflow generated by their business while other borrowers with stable employment and income are more attractive to the lenders in P2P website.

Lenders associate emotionality of reduced affective concerns with psychological distance (Lundberg et al., 1972; Nook et al., 2017; Van Boven et al., 2010). When people make decisions for a stimulus that is psychologically close, they consider more emotion-related information than when they make same decisions regarding a stimulus that lies at a greater psychological distance (Williams et al., 2014). For example, borrowers try to appeal to audiences when they list their information on the P2P platform, which can be achieve affectively. Huang et al (2020) show the borrowers' self-description that trust-related words will positively influence lenders' affective trust. In fact, through their investigation, most of descriptions evoke positive emotions. Therefore, if the psychological distance is high, the P2P borrowers' appealing will be less effective. Against this background, we hypothesize:

H1. *Psychological distancing is negatively related to funding success.*

Moderating role of Language intensity

Another theory that is related to style of words is LET. LET centers on how message features such as word choice and sentence structure, positively or negatively violate the expectation forming of the recipients (Averbeck, 2014). It also posits that the persuasion power of the message may depend on intensity of the message interacting with other factors and predicts how language intensity interacts expectation to enhance/reduce persuasion effect (Burgoon et al., 2002). Positive violations are exceeding the expectations, which will cause greater attitude change and enhance the persuasiveness of the message while negative violations means the failure to meet the expectations and result in an opposite effect.

Drawing on LET, we argue language intensity strengthen the negative relationship between psychological distancing and funding success. Bowers (1963) posits the extremities in use of words produce a moderating effect and conclude that language intensity has to be treated as a complex variable that is subject to interactions with other variables. We propose when the psychological distance is high, the high language intensity will push the distance further and result in unpleasant interpersonal outcomes. This is because high intensity is likely to negatively violate the expectation of the lenders, which causes an opposite effect. In P2P lending market, borrowers are expected to present evidence-based element objectively (Lee et al., 2019; Allison et al., 2013) rather than use high intensity language to persuade lenders. In an internet-mediated transaction environment, the participants are only allowed narrow bandwidth[12] (Jensen et al., 2013). If their narrative deviates from narrow bandwidth, this will cause expectancy violations and therefore affect lenders' decision. In an experimental analysis, Jensen et al., (2013) find online product reviews with high language intensity are more likely to cause a negative language expectancy violation, which in turn reduces credibility. The P2P platform can be regarded as an online financial product market, which shares similar features and expectancy with other online markets.

Drawing on LET, most of extant research investigates the moderating effect of language intensity on source characteristics and persuasiveness. The source characteristics the literature studied focus on gender or other demographic characteristics (Burgoon et al., 2002). There are rich characteristics that can be and should be studied in the persuasion process. When the borrowers ask for funding in P2P platform, they may want the money desperately and use high-intensity language. Lenders' expectancy may be negatively violated. Meanwhile, lenders perceive psychological distancing from the style words the borrowers choose and this can influence leaders' decision on the campaign. Both language intensity and psychological distancing are embedded in P2P borrowers' description. Given the literature suggests high language intensity are more likely to cause a negative language expectancy violation, we argue that language intensity will enhance the negative relationship. Against this background, we hypothesize:

H2. *Language intensity negatively moderately the negative relationship between psychological distancing and funding success.*

[12] Following Burgoon (1995) and Burgoon and Miller (1985), we define bandwidth a range of persuasion strategies that can be effective in senders' persuasion attempt.

3.3 Research Setting

Renrendai is one of the largest P2P platforms operating in Mainland China and many papers use Renrendai's data due to representativeness (Mi and Zhu, 2017; Tao et al, 2017; Yao et al,2019; Li and Hu , 2019; Ding et al, 2018). Renrendai was established in 2010 and registered capital is 100 million RMB. In January 2014, Renrendai acquired venture capital (130 million dollar) from TrustBridge Partners. Until Oct 2018, the accumulative transaction reaches 71.4 Billion RMB (Renrendai.com). Now it has been ranked 2nd by wdzj.com among all 1881 P2P platforms in China. The loan application process in Renrendai is as following. Frist, borrowers submit their application form with their national ID number and other personal information. They would need to specify the requested loan amount, the interest rate they would like to pay, the duration they will pay back the loan, the purpose of borrowing and any other information they find helpful to their application. The platform will then access their ID, mobile number, address, employment, income and etc. it is suggested by the platform to disclose additional information such as education qualification, car and house ownership, marital status and other professional certification to promote their credit rating. At the end, the platform will verify the information submitted by applicants and assign a credit grade to the applicant. Each credit grade corresponds different credit rating from HR (high risk) to AA (very safe). The loan application process is very similar among Chinese P2P platforms. Although some major platforms such as Paipaidai encourage applicants upload photos and declare friendship, they all use in-house credit rating to classify borrowers. Establishing in-house credit rating requires information from submitted documents.

As mentioned above, due to lack of reliable personal credit rating agency, Chinese P2P platforms rely heavily on offline authentication to reassure investors. Renren dai cooperates with Ucredit (www.ucredit.com), an offline verification company, to do physical site check.

After physical visiting to home and working address, Ucredit will then tell Renrendai the qualified borrowers. All borrowers recommended by Ucredit will be assigned A class credit rating when their application listing online. Ucredit not only acts as an offline partner of Renrendai, it but also operates as an independent Micro Finance company. They provide money mainly for small and medium enterprises (SMEs) which have financial constraints. They are experienced in dealing with loan applicants' materials. At beginning of the operating, Renrendai adopted a pure online model similar to US platforms. Yet, because of severe asymmetric information issue in China, Renrendai has no choice but to start offline services in order to alleviate the issue. The platform makes profits by offering the matching services among borrowers and lenders. According to the credit rating, borrowers need to pay a premium from 0% to 5%. Apart from that, there is a 0.3% annual service fee based on outstanding loan principal. If the listing is unsuccessful, there is no any fees for borrowers. Renrendai only does not charge originated fees from investors. The investors will be charged only if they wish to pull their money out ahead of agreement date. Although collateralized P2P model is increasingly popular in China during recent years, loans in Renrendai are all uncollateralized. The loan terms are up to 3 years and loan size is range from 3000 RMB to 500,000RMB. Renrendai guarantees the principal of investors will be paid back. To do so, the platform established a reserve fund to cover the default and the late payment. The fund is topped up constantly by the service fees charged. At the meantime, the platform will try to recover the loan by emailing, texting and calling the borrowers. If the platform fails to collect back the loan, a collection agency will step in. When the money is collected, the platform will put it into reserve fund.

3.4 Data Description

Our data is retrieved from Renrendai platform and refer to all the listings between 1 Jan 2015 and 31 Dec 2015. We only focus on the loans that will be used to fund a start-up-related activity (based on the listing title). There are 43824 listings in our dataset. After deleting missing values, our dataset of made of 33028 listings. Among them, 9020 listings are fully funded. Each listing has a set of information available to potential investors when they made their investment decisions. These information include (1) the loan characteristics such as interest rate, loan amount, loan duration, credit and financial information, (2) demographic information such as gender, age, education background, marital status, employment status, personal income and employment status and so on and (3) loan description which is borrowers' narratives. To test the borrowers' choice of offline authentication and mitigate the sample selection bias, we use Branch as dependent variable. We create Branch which is equal to one if borrowers' location has an offline branch to do physical check and zero, otherwise. This is also a proxy for financial accessibility. People who live in some Chinese major cities are better off because in these cities, they easily get access to funding for their business. We conjecture that borrowers are not willing to do offline check if there is no offline branch in their city although by doing so, loan application is more likely to be successful (Tao et al, 2017) and there is no additional cost.

Our main dependent variable is funding success. The Funded is a dummy variable, which takes one if the loan application is successful. The rationale is that if the linguistic style is able to influence lenders' decision making, this will be manifested by the probability of funding success.

The independent variables to proxy psychological distancing are the number of second person pronouns (NumYou) and the number of negations (Negations)[13]. Toma and Hancock (2012)

[13] Examples of second person pronouns and negations are as followings: "The purpose of the loan is to buy machinery and pay deposit of the business. As long as the business works well, I can have approximately RMB 20,000 per month. There is no problem to pay back loan. Hope you trust me." "I have some spare time after work so I would like to apply for my first loan here to do some small business. I have no idea whether it is trustworthy or not trustworthy (NT: in English, we may omit last

suggest linguistically, psychological distancing manifest itself by a decrease in first person pronouns and an increase in negations. Pronouns can reveal how people are referencing those outside of or in the interaction (Tausczik and Pennebaker, 2010).The use of first-person plural pronouns ("we" and "ours") implies inclusiveness while second and third- person pronouns ("you" and "they") distance them away from the interaction. Moreover, in the Chinese linguistics, the use of second person pronouns "ni" is associated with perceived disrespect (Jiang et al., 2013). As it is less likely to use first and third-plural pronouns when writing to an unknown audience (Toma and Hancock, 2012), we follow Toma and Hancock (2012) to measure psychological distancing using the number of second person pronouns (NumYou) and in light of extant research (Simmons et al, 2005; Revenstorf et al, 1984), we predict the use of "you" would lead to unsatisfactory funding outcome.

Toma and Hancock (2012) also suggest psychological distance increases along with use of negations. Xu (2015) invited 113 Chinese college students with average age of 18.8 year old to participate in an experiment related to negations communication. The results show that senders' negations communication has a negative effect on receivers' impression. To be specific, when the positive information is sent with negations, receivers tend to response negatively. Moreover, using a doctor-patient experiment, Burgers et al (2012) find patients feel more negative when positively framed bad news with negations are delivered than with affirmations. We therefore predict a negative relationship between negations and funding outcome.

The moderator is language intensity. Li and Zhan (2011) suggest number of exclamation marks, strong positive emotions and strong negative emotions can be used to identify language

---

trustworthy but in Chinese, we keep it) so I just try a small amount of money. I am working at Changsha and have bought a 150 square feet flat in my hometown."

intensity. We therefore first follow Huang et al. (2020) to calculate the sentiment index of borrowers' description. Then we follow Li and Zhan (2011) to create two dummy variables, extremPos and extremNeg. extremPos takes one if the sentiment index is higher than 90% percentile, or zero otherwise. extremNeg takes one if the sentiment index is lower than 10% percentile, or zero otherwise. Following Li and Zhan (2011) and Han et al (2018), last variable to proxy language intensity is the number of exclamation mark (Exclamation).

Control variables include number of words used to describe the project, education, age, loan amount, loan duration and monthly income. A longer text is considered to be a signal of openness and transparency from the borrowers' end of the transaction and therefore we expect that lenders are willing to fund loans that are well described and explained (Dorfleitner et al, 2016). In addition, the longer text, the more likely positive words could be included. Yet, if the description text is too wordy, the investors may not have enough patience to go through all the text. Since they only invest a small amount of money, it is not worth to cost much time (Dorfleitner et al, 2016). Taken together, we expect a positive association between the number of words and investors' action. Income indicates income range where n=0(less than 1000RMB), 1 (between 1000 and 2000RMB), 2(between 2001 and 5000 RMB),3 (between 5001 and 10000 RMB), 4(between 10001 and 20000 RMB), 5(between 20001 and 50000 RMB), and 6 (over 50000 RMB); Degree indicates the education level where n=0 (high school or lower), 1(junior college), 2(undergraduate), and 3 (master or above); NonOffline is coded one if the listing has not been verified through the offline authentication. Credit rating is provided by Renrendai, which from HR to AA. We code 0 for HR, 1 for E, 2 for D, 3 for C, 4 for B, 5 for A and 6 for AA.

3.5 Research Design and Model Specification

Our research interest is whether psychological distancing measured by linguistic styles is associated with investors' decision making. Since we cannot observe language styles from the borrowers who use offline authentication service and the offline branch staff will upload a template to description if they use these service, we need to confine our sample only to the listings without offline authentication. Yet, borrowers who do not choose authentication service may follow specific demographic pattern so simply regressing this sample is not appropriate. It is possibly that factors influencing borrowers' choice of authentication service could be also correlated with lenders' decision. In this case, the coefficients of linguistic styles would be correlated to error term. There is a selection bias if we simply use ordinary least squares (OLS) or generalized least squares (GLS) (Wang et al, 2008).

In order to overcome the selection bias, we adopt Heckman's (1979) seminal work known as two step method. In short, Heckman two step method corrects sample selection by first using a probit model to regress the individual characteristics on offline authentication decision. The outcome variable is regressed in the second step using least square on the independent variables which are our interests and fitted value from the first step (selection equation).

The identification of Heckman method needs a valid exclusion restriction in the selection equation. To do so, we need a variable that is in the selection equation but excluded in second stage equation. If we fail to meet exclusion restriction in the selection equation, the estimates of second equation are likely to be biased (Angrist & Krueger, 2001; Hamilton & Nickerson, 2003). In this study, we use Branch. Branch indicates whether the borrower is able to find an offline branch to do the physical check. This variable is valid because in the second step, Branch is not related to the dependent variable, funding success. Hence, we exclude Branch in our main analysis. If there is no local branch, borrowers may have to apply online without authentication. Other variables include Mortgage, car loan, age, income, education and gender. We conjecture that whether borrower choosing offline authentication or not is associated with their self-confidence. If borrowers have previous loan records (mortgage or car loan), they may be willing to do offline authentication because loan history may signal good credit trustworthiness. Also, female borrowers are more likely to use offline service because they are more conservative compared to male borrowers who are bold and risk taking. We use STATA command, Heckprobit, to estimate two steps mentioned.

$$\mathrm{Prob(NonOffline} = 1) = \beta_0 + \beta_1 \mathrm{Mortgage} + \beta_2 Carloan + \beta_3 Branch + \beta_4 age + \beta_5 Income + \beta_6 Education + \beta_7 Gender + \mu \qquad (1)$$

$$\mathrm{Prob(Funded} = 1) = \beta_0 + \beta_1 Linguistic\ styles + \beta_2 CreditRating + \beta_3 Interest + \beta_4 Duration + \beta_5 Ln\ Amount + \beta_6 age + \beta_7 Income + \beta_8 Education + \beta_9 Gender + \mu$$
(2)

Where (2) is the main equation, while (1) is the sample selection equation. Linguistic styles includes *NumYou* and *Negations*.

3.6 Empirical Results

Table 9 shows the descriptive statistics. Panel A is for unsuccessful loans and Panel B displays statistics of successful campaign. Compared to successful loans, unsuccessful loans have, on average, higher numbers of NumYou (0.02 against 0.00) and Negations (0.360 against 0.140). These statistics provide initial support to our hypotheses. Table 10 shows the VIF test results. The VIF test is to assess multicollinearity. The maximum VIF is 6.39 and the mean VIF is 2.19. Given the threshold of is 10, there is no sign of multicollinearity.

Table 9. Descriptive Statistics

Panel A Unsuccessful Loans					
variable	N	Mean	sd	min	max
WordCount	22125	58.03	52	4	718
CreditRating	22125	0.0300	0.260	0	5
LnAmount	22125	10.88	1	8.010	13.12
Income	22125	3.060	1.170	0	6
interest	22125	12.65	0.670	8	13
months	22125	20.20	6.180	3	36
age	22125	30.41	6.240	22	56
marriage	22125	0.620	0.570	0	3
gender	22125	0.160	0.360	0	1
NumYou	22125	0.02	0.170	0	5
Negation	22125	0.360	0.770	0	12
extrmPos	22125	0.0900	0.280	0	1
extrmNeg	22125	0.0300	0.180	0	1
Exclamation	22125	0.210	0.740	0	40
NonOffline	22125	1	0.0400	0	1
mortgage	22125	0.120	0.320	0	1
carloan	22125	0.0700	0.250	0	1
Branch	22125	0.430	0.500	0	1
Panel B Successful Loans					
variable	N	Mean	sd	min	max

WordCount	8924	117.2	57.84	19	492
CreditRating	8924	4.440	1.470	0	5
LnAmount	8924	10.82	0.650	8.010	12.44
Income	8924	3.610	1.310	0	6
interest	8924	11.43	0.960	8	13.20
months	8924	25.26	9.970	3	48
age	8924	37.22	8.430	20	62
marriage	8924	0.930	0.510	0	3
gender	8924	0.260	0.440	0	1
NumYou	8924	0.00	0.0300	0	1
Negation	8924	0.140	0.430	0	8
extrmPos	8924	0.140	0.340	0	1
extrmNeg	8924	0.280	0.450	0	1
Exclamation	8924	0.0300	0.290	0	11
NonOffline	8924	0.280	0.450	0	1
mortgage	8924	0.350	0.480	0	1
carloan	8924	0.0900	0.290	0	1
Branch	8924	0.870	0.330	0	1

Table 10. VIF

Variable	VIF	1/VIF
CreditRating	6.39	0.156447
interest	5.47	0.182912
months	4.44	0.225299

Lnwordcount	1.94	0.516245
age	1.55	0.647053
LnAmount	1.44	0.696090
Income	1.37	0.732255
marriage	1.32	0.758908
extrmNeg	1.27	0.787308
Negation	1.25	0.800785
extrmPos	1.15	0.866839
Exclamation	1.04	0.963718
gender	1.03	0.973436
NumYou	1.02	0.976741
Mean VIF	2.19	

Table 11. Correlation Coefficient

	Funded	WordCount	CreditRating	LnAmount	Income	interest	months	age	ma
Funded	1								
WordCount	0.446***	1							
CreditRating	0.925***	0.481***	1						
LnAmount	-0.029***	0.084***	0.061***	1					
Income	0.201***	0.199***	0.211***	0.386***	1				
interest	-0.586***	-0.366***	-0.588***	0.173***	-0.219***	1			
months	0.293***	0.077***	0.373***	0.307***	-0.044***	0.418***	1		
age	0.406***	0.241***	0.431***	0.217***	0.253***	-0.230***	0.188***	1	
marriage	0.248***	0.132***	0.263***	0.161***	0.175***	-0.117***	0.149***	0.474***	1
gender	0.122***	0.070***	0.135***	0.032***	0.013**	-0.056***	0.087***	0.058***	0.0

NumYou	-0.050***	0.062***	-0.049***	-0.024***	-0.028***	0.022***	-0.024***	-0.018***	-0
Negation	-0.142***	0.284***	-0.152***	-0.050***	-0.063***	0.062***	-0.098***	-0.097***	-0.0
extremePos	0.075***	0.342***	0.082***	0.013**	0.054***	-0.124***	-0.077***	0.036***	-0.0
extremeNeg	0.363***	0.214***	0.394***	0.056***	0.026***	-0.158***	0.261***	0.147***	0.1
Exclamation	-0.124***	0.020***	-0.136***	-0.026***	-0.016***	0.081***	-0.053***	-0.072***	-0.0
NonOffline	-0.797***	-0.243***	-0.871***	-0.097***	-0.132***	0.404***	-0.463***	-0.382***	-0.2
mortgage	0.277***	0.102***	0.291***	0.179***	0.118***	-0.026***	0.310***	0.189***	0.1
carloan	0.048***	0.00400	0.046***	0.095***	0.138***	-0.00400	0.021***	0.028***	0.0
Branch	0.403***	0.172***	0.435***	0.036***	0.118***	-0.234***	0.187***	0.154***	0.0

Table 12 shows the results of first step Heckman selection model, which is a probit regression of not doing offline authentication against factors considered to influence whether a borrower chooses to engage in offline authentication. The dependent variable, *NonOffline,* is a binary variable indicating whether a borrower engaged in offline authentication. *NonOffline* takes one if the borrowers didn't choose offline authentication, because in second step, we only focus the borrowers without offline authentication. In Table 12 model (1), we do not include the *age, marriage, income and gender*. After adding these in model (2), the R square increases to 0.437 from 0.334. The increase in R square shows high explanatory power of branch and necessity of including this variable. The value of the Chi square change (-3266.0648, $p<0.01$) at the bottom of Table 12 again implies that Model 2 has a better fit. Hence, we use model (2) in the main analysis. In line with our prediction, financial accessibility plays a key role in offline authentication. The negative and significant coefficient of branch indicates if there is a local branch in the city, the borrower is more likely to undergo offline authentication. Being able to take mortgage and car loan and receiving higher income are signals of creditworthiness and give confidence to borrowers. So the coefficient of mortgage and car loan, and income are negative and significant, which implies borrowers with mortgage and car loan, and higher income are more likely to use offline services. The significantly negative coefficient of age and negative coefficient of gender indicate younger and male borrowers are more risk-taking so they are more willing to apply online directly without offline authentication.

Table 12. Selection Equation

	(1)	(2)
	NonOffline	NonOffline

mortgage	-0.987***	-0.835***
	(-43.05)	(-34.16)
carloan	-0.262***	-0.294***
	(-7.58)	(-7.92)
Branch	-2.723***	-2.862***
	(-36.66)	(-37.70)
age		-0.0611***
		(-39.48)
marriage		-0.309***
		(-14.14)
Income		0.0206*
		(2.38)
gender		-0.360***
		(-13.86)
_cons	3.333***	5.773***
	(44.76)	(60.52)
N	31049	31049
Pseudo R-square	0.334	0.437

Log Likelihood (LL)	-10542.435	-8909.4026
Deviance (-2LL or Chi square change vs. Model 1)		-3266.0648***

*** 1% ** 5% * 10%

Z statistics in parentheses;

The empirical results of main analysis (second step) are shown in Table 13 and Table 14. Generally speaking, the use of the Heckman model is suitable because the P-Value of the Wald test from *Heckprobit* is less than 0.10 for all the model specifications. The dependent variable is funding success. The Table 13 shows the results of the number of "you" (NumYou) and the model 1 shows the coefficient of NumYou is -0.487 and significant at 1%. We then gradually add interaction terms of NumYou and language intensity variables to test moderating effects of language intensity. The table 13 model 2 shows PosYou, the interaction of extremPos and NumYou, is negative and significant (β= -3.664, p<0.01).The interaction term of extremNeg and NumYou, NegYou, is not statistically significant (β=0.082, p>0.1), while the interaction of Exclamation and NumYou, ExclamYou, is negative and significant (β=-3.801, p<0.01).

The Table 14 shows the results of *Negations*. The model 1 shows Negations are negatively related to the funding success and it is statistically significant at 1% (β=-0.097, p<0.01). Again, we put interactions of language intensity variables and Negations to the model. The three interactions of language intensity and Negations, PosNeg, NegNeg and ExclamNeg are not significant (β=-0.094, p>0.1; β=0.029, p>0.1; β= 0.018, p>0.1 respectively).

Table 13. Main results: the number of you

	(1)	(2)	(3)	(4)
	Funded	Funded	Funded	Funded
LnUSTword	0. 712***	0.706***	0.727***	0. 716 ***
	(0.235)	(0.236)	(0.237)	(0. 235)
LnUSTword2	-0.078**	-0.077**	-0.081**	-0.079**
	(0.032)	(0.033)	(0.033)	(0.032)
CreditRating	1.064***	1.069***	1.066***	1.064***
	(0.051)	(0.051)	(0.051)	(0.051)
LnAmount	-0.754***	-0.754***	-0.754***	-0.755***
	(0.020)	(0.020)	(0.020)	(0.020)
Income	0.271***	0.270***	0.271***	0.271***
	(0.018)	(0.018)	(0.018)	(0.018)
interest	0.198***	0.194***	0.197***	0.199***
	(0.070)	(0.070)	(0.070)	(0.070)
months	-0.039***	-0.040***	-0.039***	-0.039***
	(0.008)	(0.008)	(0.008)	(0.008)
age	0.023***	0.023***	0.023***	0.023***
	(0.003)	(0.003)	(0.003)	(0.003)

marriage	0.089***	0.088***	0.088***	0.090***
	(0.034)	(0.034)	(0.034)	(0.034)
gender	0.027	0.028	0.028	0.026
	(0.051)	(0.051)	(0.051)	(0.051)
NumYou	-0.462***	-0.439***	-0.512**	-0.415***
	(0.165)	(0.163)	(0.223)	(0.157)
extremPos		-0.073		
		(0.078)		
PosYou		-3.673***		
		(0.245)		
extrmNeg			0.072	
			(0.092)	
NegYou			0.121	
			(0.274)	
Exclamation				-0.018
				(0.024)
exclamYou				-3.855***
				(0.211)
cons	1.058	1.097	1.046	1.044

	(0.870)	(0.873)	(0.873)	(0.871)
/athrho	0.098*	0.098*	0.098*	0.098*
	(0.051)	(0.051)	(0.051)	(0.051)
N	31049	31049	31049	31049

Standard errors in parentheses; * p<0.1, ** p<0.05, *** p<0.01

Table 14. Main results: the number of negations

	(1)	(2)	(3)	(4)
	Funded	Funded	Funded	Funded
LnUSTword	0. 658***	0. 628**	0.689***	0.677***
	(0. 242)	(0. 246)	(0.248)	(0. 243)
LnUSTword2	-0.067**	-0.062*	-0.072**	-0.069**
	(0.034)	(0. 035)	(0.035)	(0.034)
CreditRating	1.064***	1.071***	1.066***	1.064***
	(0.051)	(0.020)	(0.051)	(0.051)
LnAmount	-0.752***	-0.751***	-0.751***	-0.753***
	(0.020)	(0.020)	(0.020)	(0.020)
Income	0.267***	0.266***	0.266***	0.267***
	(0.018)	(0.018)	(0.018)	(0.018)
interest	0.203***	0.198***	0.202***	0.202***

	(0.070)	(0.069)	(0.070)	(0.070)
months	-0.040***	-0.040***	-0.041***	-0.040***
	(0.008)	(0.008)	(0.008)	(0.008)
age	0.023***	0.023***	0.023***	0.023***
	(0.003)	(0.003)	(0.003)	(0.003)
marriage	0.087**	0.085**	0.087**	0.087**
	(0.034)	(0.034)	(0.034)	(0.034)
gender	0.029	0.031	0.030	0.028
	(0.051)	(0.051)	(0.051)	(0.051)
Negation	-0.081***	-0.072**	-0.089***	-0.088***
	(0.028)	(0.028)	(0.030)	(0.029)
extrmPos		-0.037		
		(0.080)		
PosNeg		-0.077		
		(0.088)		
extrmNeg			0.025	
			(0.111)	
NegNeg			0.042	
			(0.060)	

Exclamation				-0.030
				(0.030)
exclamNeg				0.020
				(0.014)
cons	1.086	1.174	1.037	1.076
	(0.870)	(0.873)	(0.879)	(0.872)
/athrho	0.099*	0.100*	0.100*	0.099*
	(0.051)	(0.051)	(0.051)	(0.051)
N	31049	31049	31049	31049

Standard errors in parentheses; * p<0.1, ** p<0.05, *** p<0.01

3.7 Discussion and Conclusion

3.7.1 Contribution

The intent of this research is to advance of our knowledge of the effect of psychological distancing on P2P funding performance. We bridge the P2P lending literature and psycholinguistics literature and set out to explain how psychological distancing manifested by linguistic styles can influence lenders' decision on P2P funding campaign. Building upon CLT, we find psychological distancing is a key predictor of P2P funding success for entrepreneurs after controlling for sample selection bias. Our results are consistent with Parhankangas and Renko (2017) which suggests avoiding psychological distancing is important for social entrepreneurs in reward-based crowdfunding campaign. We extend the their results to debt-based crowdfunding. This finding is also consistent with psycholinguistics literature which suggests that psychological distancing is associated with negative interpersonal outcome (Simmons et al, 2005; Revenstorf et al, 1984). Specifically, the number of "you" and the number of negations used in borrowers' description are negatively related to the willingness of the lender to support the funding campaign. Moreover, drawing on LET and extant literature which suggests a moderating role of language intensity (Bowers, 1963; Burgoon et al., 2002), we posit the language intensity tends to strengthen the negative relationship between psychological distancing and funding success. Our empirical results provide general support for the argument.

Our contributions to the literature are twofold. First, this paper contributes to the role of language in crowdfunding (Parhankangas and Renko, 2017; Anglin et al., 2018; Herzenstein et al., 2011; Allison et al., 2013; Majumdar and Bose, 2018). This strand of literature tends to find how borrowers use language strategically to improve the likelihood of funding success in different types of campaign. Studies rely on message forming or signaling theory and investigate the narratives that have been written on purpose (Huang et al., 2020; Anglin et al., 2018; Block et al., 2018). Yet, borrowers may unconsciously write something that influence lenders' decision. Drawing on CLT, we first argue the high psychological distancing produced by unconscious use of linguistic style will have reduced affective concern, which has a negative influence on P2P funding performance.

Second, we contribute to LET by exploring the moderating role of language intensity in P2P lending. Prior studies find a detrimental role language intensity in internet-mediated environments (Han et al., 2018; Li and Zhan, 2011). However, their study does not find support LET which posits a moderating role of language intensity interacting with source characteristics (Han et al., 2018). We show the moderating effect of language intensity with psychological distancing in determining lenders' perception and find support for LET. In addition, existing researches examine psychological distancing and language intensity separately. However, language intensity that is closely related to tone of communication cannot be separate during the communication. Recipients perceive the quality of messages with both linguistic styles and tone of communication together. Therefore, investigating how linguistic styles interact with language intensity is necessary. Our results suggest that language intensity negatively strengthen the negative association between the number of second person pronouns and funding success but has no significant effect on the relationship between negations and funding success.

3.7.2 Implication

Our results also have some practical implication. For borrowers, it is important to avoid psychological distancing with lenders through linguistic styles and tone of communication. For example, borrowers should reduce the use of second person pronouns and negations. Especially in China, the use of second person pronouns, “ni” shows disrespectful. Recipients are confused if high frequency use of negations in the communication because it is difficult to process negations for human beings.

This study is not without drawbacks. First, although we construct our variables based on loan description, we did not pay attention to the specific content they disclosure. Yet, it is believed that investors also will consider the information they mentioned. So researchers are encouraged to analyze the specific content when doing further investigation. Second, this research is solely based on the data from one Chinese crowdfunding platform, Renrendai. Albeit Renrendai is one of the biggest and the most popular platforms in China, it is unique among other platforms in many aspects. This may limit the generality of our study. Hence, it is suggested future studies may explore different platforms to validate our results.

Chapter 4 The effect of Peer-to-Peer lending on financial exclusion: Evidence from China

Abstracts

The association of financial technology (fintech) and financial exclusion has attracted attention since rapid growth of fintech innovation. This study investigates the funding probability of the financial excluded borrower in a large P2P lending platform. Using loan-level data from a lending Chinese P2P company, we find there is a negative indirect effect of financial exclusion on funding success through credit score. In a moderated mediation analysis, we also find new business model such as offline authentication and education qualification positively moderates the linkage between the financial excluded and credit score and therefore negative indirect effect of financial exclusion on funding success is overturned when the excluded borrower has conducted offline authentication and obtained higher education qualification. Lastly, we examine the determinants of offline authentication decision. We find the borrowers in a city with better financial infrastructure are more willing to conduct authentication. However, the financial excluded borrowers are less likely to conduct offline authentication.

"We envision a financially inclusive world where all people hold the power to improve their lives. More than 1.7 billion people around the world are unbanked and can't access the financial services they need. Kiva is an international nonprofit, founded in 2005 in San Francisco, with a mission to expand financial access to help underserved communities thrive. We do this by crowdfunding loans and unlocking capital for the underserved, improving the quality and cost of financial services, and addressing the underlying barriers to financial access around the world. Through Kiva's work, students can pay for tuition, women can start businesses, farmers are able to invest in equipment and families can afford needed emergency care."

——Kivo.org

4.1 Introduction

The development of Financial Technologies ("Fintechs) has changed the financial world (Chen and Bellavitis.,2020; Palmie.,2020; Panos and Wilson., 2020). It can offer new and fast financial services to a larger group of people by integrating finance and technology. The Fintechs are expected to replace/complement traditional financial systems and adopt new business model benefiting from new technologies and market conditions. For example, with smartphone, people can conduct most of personal banking business which can only be done face-to-face previously. Moreover, one of the Fintechs, Peer-to-Peer lending (thereafter, P2P lending), enables borrowers to get unsecured loans (e.g., loans without collateral) from individual lenders in a P2P platform (Lin et al., 2013). The unsecured loans lift the barriers of people getting the loan to some extent so researchers consider P2P lending may allay financial exclusion (Sparreboom and Duflos, 2012).

Financial exclusion has been regarded as a multi-dimensional concept that consists of a set of obstacles to accessing and using basic financial services (Kempson and Whyley, 1999a,b; Devlin., 2005). To alleviate financial exclusion is a critical element for poverty reduction and social development. It is expected to result in greater financial stability and economic growth[14]. It is important to distinguish two seemingly similar words, access and use, in the context of financial exclusion. Access refers to the supply of financial products while use is determined by both supply and demand (World Bank, 2014). This study focuses on examining the credit provision perspective of financial exclusion. This is a supply-side/access issue coming from financial institutions unwilling to provide credit to certain group of people. Kempson and

[14] https://www.adb.org/publications/financial-inclusion-asia-overview

Whyley (1999a,b) have identified five dimensions of financial exclusion: (1) access exclusion coming from adverse risk assessment or branch closures; (2) condition exclusion that means individuals are excluded because financial products are not designed for them; (3) price exclusion where individuals are not able to afford financial services; (4) marketing exclusion where individuals are excluded by targeting marketing and sales; (5) self-exclusion where individuals choose not to use financial products due to preferences, cultural norms, religious norms and etc. The first four dimensions of financial exclusion are related to supply-side barriers while the last one coming from the demand side. Actually, financial exclusion can be assessed in many ways such as money transmission, credit, pensions, insurance cover, financial education, debt and etc (McKillop and Wilson., 2007). Ibtissem and Bouri (2013) suggest low income households are generally considered that they are financially excluded. In line with Ibtissem and Bouri (2013), this study uses low income to proxy the financial exclusion. Due to only a few people who are able to access to basic financial services nowadays, alleviating financial exclusion is yet an unfinished task. The obstacles that prevent people access to basic financial services include physical, bureaucratic, financial and trust barriers (Aggarwal and Klapper, 2013). Innovation in fintech opens a potential avenue for addressing the problems. For example, In a case study of Autazes, a county located in Amazon region, Diniz et al (2012) find information and communication technology-based (ICT) branchless banking enables tens of millions of low-income local residents access to financial services; there is no other way to access to banking services for most of them. In the descriptive analysis of internet finance disparity among Chinese provinces, Hasan et al (2020) suggest in the eastern and the southern of China, people have largely adopted Fintechs such as internet and mobile banking and this leads to alleviating financial exclusion in these provinces. In addition, Inception of P2P lending will make more people get access to credit and have a deeper reach than banks because it can connect potential lenders and borrowers directly without financial intermediaries (Sparreboom and Duflos, 2012; Komarova Loureiro and Gonzalez., 2015). Using an US survey data asking people how they finance their business, Schweitzer and Barkley (2017) find borrowers who have been denied credit by the bank turn to online lenders to provide credit for their business; they conclude online lenders can provide credit to the borrowers who are not qualified for bank financing. The physical barriers such as no bank branches can be also addressed by P2P platforms as the transactions can be completed online. Yet, the some negative effects found in fintech lending poses a question on fintech's prospect for mitigating financial exclusion. Based on Prosper' data[15], Freedman and Jin (2011) find P2P lenders are more willing to serve consumers who can traditionally obtain funds from banks as the platform tends to exclude more and more subprime borrowers. In this case, P2P platform will become an alternative choice of privileged borrowers rather than a channel for financially excluded borrowers. Lin and Viswanathan (2015) find evidence that P2P lender are more willing to fund local borrowers and suggest internet technology is less likely to address home bias, which means borrowers from rural and remote area may not benefit from fintech innovation. Fuster et al. (2019) find although fintech lenders has increased their market share in mortgage lending from 2010 to 2016 from 2% to 8%, fintech borrowing is positively associated with bank branch density and they don't target borrowers who are unable to get access to traditional finance. This again implies again fintech cannot enable credit to a larger group of people. In addition, regulatory and supervisory authority may not be able to response to the new technology timely so fintech players could not fit into current regulatory regime and their rights may be compromised. Concerns of cyber security are another issue to discourage participation of new technology.

Research Questions

Among all the concerns mentioned above, the asymmetric information is the main concern of lenders in emerging market to participate in P2P market (Chen et al., 2018). Asymmetric

[15] Prosper is a leading P2P company in the US.

information leads to capital rationing of the lenders (i.e., each lenders only contribute to small amount of funding) and, as a result, restraints the participants' ability to raise the funds (Matson., 1999). If there is no lenders willing to supply credit, no one can benefit from the fintech innovation. The P2P platform therefore have created in-house credit rating system to tackle asymmetric information in the country without nationwide-recognized credit rating system such as China. After submission of documents that the platform asks, the platform will assign a credit score to borrowers. The credit system plays an important role in encouraging participation. Evidences show that lenders are more willing to bid in higher rating campaign and borrowers with higher credit rating are more like to be funded (Chen et al., 2018; Tao et al., 2017). However, the financial exclusion is directly related to low income, so the financial excluded may not be able to have a high credit score and as a consequence, they are less likely to have their loan requests funded. To date, the study to examine the interplay between the financial excluded and P2P lending success is scarce and contradictory (Komarova Loureiro and Gonzalez., 2015; Lin and Viswanathan, 2015). More importantly, little is known about the role of credit score in the linkage between the financial exclusion and P2P funding success. This is probably because credit score is not assigned by the platform in developed countries (Chen and Han, 2012). The research gap is problematic given the alleviation of asymmetric information and funding performance of financial excluded borrowers rely on credit score system (Chen and Han, 2012). The credit score may be the key to allay financial exclusion in P2P lending market. Therefore, we raise the first research question: Does the credit score play a mediating role in the association between the financial excluded and funding success? The P2P platform in China offers a great laboratory. The number of P2P borrowers and lenders in China have increased significantly in recent years and the credit score are assigned by Chinese P2P platforms (Huang et al., 2020). The mediation hypothesis is consistent with data generation process given the fact that the borrower submits their application first, then is assigned a credit score and finally is accessed by the lenders. Empirical evidences also find credit rating has a direct impact on funding success (Tao et al., 2017; Chen et al., 2018). Moreover, low income households are financially excluded because of their low credit score assigned by the credit bureau (Bridges and Disney, 2004). We therefore conjecture that the financial excluded has an influence on funding success mainly through the channel of credit score. In addition, if the mediating effect of credit score indeed exists, this implies the P2P platform has an influence on both the financial excluded and lenders' decision making as the credit score is provided by the platform.

To further reassure participants, major P2P platforms have adopt offline authentication to mitigate lenders' concerns about information asymmetry and as a risk reduction tool for platforms themselves. Borrowers who have been offline authenticated can positively signal their creditworthiness to the investors. Using data from Renrendai, Tao et al (2017) find that offline authentication can significantly reduce the problem of asymmetric information. They find the listings with offline authentication are more likely to get funded, decrease the interest rate and repay the loan on time. In an analysis of Canadian platforms, Cumming et al (2019) show platform due diligence such as site visit has a positive impact on funding success, number of contributor and amount of capital raised. However, due to the uniqueness of offline authentication, many P2P research excludes loans with offline authentication from their data sets and claims these are not typical P2P loans (Chen et al, 2019). Application of such business model in the country with severe asymmetric information may promote fintech innovation's potential of addressing the problem of financial exclusion. Moreover, level of education plays a key role in financial exclusion (McKillop and Wilson, 2007; Kobu, 2015; Solo, 2008; Amaeshi, 2006). Amaeshi (2006) attributes financial exclusion in Nigeria to illiteracy. Providing a higher education qualification also signals creditworthiness of the borrowers since they are expected to have a higher income (Xu et al., 2018). Therefore, the financial excluded with a higher educational attainment may enable lenders to trust them due to their ability to make money and repay the loan (Swaidan et al., 2003). As the one of main issues that prevent lenders' participation is asymmetric information, we expect the funding performance of the financial excluded will be better after effective signals (e.g. offline authentication and education qualification) have been provided. So far, P2P platforms' offline check model, educational attainment and their impact on the association between P2P lending success and financial exclusion has not been sufficiently investigated. Although studies have showed offline authentication and educational attainment play a positive role in funding success for borrowers from a wide spectrum of the population (Li and Hu., 2019; Tao et al., 2017), whether the conclusion still holds for the financial excluded remains unknown. Without such information, related parties such as P2P platforms and the government cannot act accordingly to reduce financial exclusion. Given the fact that the borrower will be assigned credit score after offline authentication and submission of education information, to investigate how offline authentication and education background moderate the financial excluded and credit score and consequently influence funding performance is more appropriate. This is so called moderated mediating effect of offline authentication and education attainment. Hence, the second research question is: Does offline authentication and education attainment positively moderate the financial excluded and credit score and improve the probability of funding success? Lastly, the empirical evidences mentioned above suggest a positive role of offline authentication on funding performance. We will try to answer are the financial excluded more likely or less likely to conduct offline authentication.

Using a large Chinese P2P platform data, we find credit score negatively mediates the financial excluded and funding success. We further show that the negative effect of the financial excluded on funding success is mainly through the channel of credit score which accounts for about 98% of total negative effect. This study contributes to previous P2P and financial exclusion literature by documenting the mediation effect of credit score plays an important role in the relation between the financial excluded and funding performance. The moderated mediation analyses show that both educational attainment and offline authentication have a positive influence on the association between the financial excluded and funding success through credit score. The negative effect is overturned when the financial excluded have obtained a higher education qualification or conducted offline authentication. The finding suggests that P2P lending doesn’t necessarily reduce financial exclusion but it can achieve the goal though education or by some new business model innovations such as offline authentication.

4.2 Literature Review and Hypothesis Development

Many financial services such as supply of short term credit are regarded as essential services[16]. Rejection from such activities can be considered as financial exclusion[17] (Howell and Wilson, 2005). Ozili (2020) suggests some special agents such as micro-finance institutions (MFIs) and post offices can make excluded people access to financial services better than other traditional financial institutions do. For example, postal operator has a unique network of counters so they are able to deliver their financial services to the excluded people facing physical barriers, especially in rural areas of developing countries (D'alcantara and Gautier, 2013). Ozili (2020) suggests to combat financial exclusion, ideally, the special agent should: 1) have specialities in this area, 2) understand the characteristics of the excluded, 3) address current issues for improvement through innovation and 4) bring the excluded into the formal financial system so they can have access to the formal financial services and product. The P2P company has

[16] See eg Office of Fair Trading (UK), Vulnerable Consumers and Financial Services: the Report of the Director-General’s Inquiry OFT255 (1999) 19.
[17] As this study mainly talks about credit provision perspective of financial exclusion, the following hypotheses are all related to loan application of the financial excluded.

specialities in helping borrowers to match lenders (Yan et al., 2018; Berger and Gleisner., 2009). The borrowers they deal with are normally rejected credit by formal financial institutions (Hou et al., 2019). The company knows the borrowers very well as they have a large amount of borrowers' information and try to match the lenders through self-innovative design (Mi and Zhu, 2017; Tao et al, 2017). These make P2P company an another suited special agent for alleviating financial exclusion.

The poor and other disadvantaged groups such as low income households are most likely to be excluded. Based on the agency theory, the exclusion of disadvantaged people from traditional financial institution is attributed to the high level of information asymmetry that will cause the additional cost to issues of screening, monitoring and enforcement (Ibtissem and Bouri, 2013). Taking loan application for example, consumers' capability to obtain credit is according to the resources at their command while the low income borrowers are short of resources, thereby lacking of credit qualification (Zhu and Meeks, 1994). Similar to Zhu and Meeks (1994), Barakova et al (2013) suggest high income Households are able to secure mortgage financing but low income people without stable employment and collateral find difficult to obtain loans. Financial vulnerability and exclusion are highly associated with low income (OFT 1999). In a theoretical analysis of Chinese famer data, Tan (2014) estimates about 80% of low and medium income famers are not able to have adequate credit because providing small loans to famers is not profitable for banks. Given the discussion above and the P2P lending market we focus on, we use low income borrowers to proxy the financial excluded.

Moreover, the credit score system may make the low income households excluded even further. The Fair Isaac Corporation (FICO) in the US created a formula to calculate individual credit score that has been used by major credit reporting agencies. Yet, the algorithm is still a secret, although people believe it is directly related to the income (Arya et al., 2013). In the UK, due to adverse scoring from credit bureaux, low-income households are financially excluded from the prime credit outlets (Bridges and Disney, 2004). The low-income borrower in general is more likely to be given a low credit score as the risks caused by asymmetric information (e.g., select a less creditworthy borrower from high creditworthy borrowers) are high. However, building a credible credit scoring system requires huge input at beginning. Not every company can afford the cost. In the context of Chinese P2P lending market, there is no nationwide credit rating system for individuals so Chinese P2P firms use their in-house credit score to classify high or low risk borrowers. It is common that a borrower obtains different credit score in different P2P platform (Zhao et al., 2016). Their credit system is opaque. Some borrowers who upload many optional certificates still get a lower rating while some borrowers only submit compulsory documents but get the highest rating (Zhao et al., 2016). Yet, recent developments on big data and credit scoring may help platforms build a self-learning system that is producing more reliable results as time goes on when big data on borrowers helps improve the system (e.g., See ONAY and ÖZTÜRK(2018) for a review). Hence, most of lenders rely on the credit rating to make decisions as the most efficient way for lenders to make decisions is to look credit rating, which lowers their screening and transaction cost. Moreover, P2P lenders usually need to make quick decisions (Liao et al., 2019). The only way to achieve that is again to look at credit rating that offers the most informative signal. Low credit rating given to the financial excluded borrowers fails to signal their creditworthiness and are therefore less likely to succeed in funding campaign. This is because P2P lenders might think high information asymmetry is associated with the financial excluded. We argue that the adverse scoring to the financial excluded reduces the probability of financially excluded borrowers being funded. Against this background, we hypothesize:

H1: *The credit score mediates the relationship between the financial excluded and funding success.*

Face to face interaction with loan officer.

To overcome the difficulty of verifying documents and alleviate asymmetric information, offline authentication mechanism has been created and applied by some Chinese P2P platforms. Offline authentication includes physical site visits to verify borrowers' income statement, working place and other key documents. The offline investigation mitigates asymmetric information and assure lenders by auditing the documents borrowers provide as fabricating documents are common online (Toma and Hancock, 2012). Offline loan officers will have a face-to-face interaction with the borrower and they may reject borrowers' loan application if they find borrowers were faking the documents or unqualified. In an analysis of offline authentication loans in a Chinese P2P lending platform, Tao et al. (2017) show borrows who have conducted offline authentication will be awarded higher credit score and find the effect of gender, education and marital status on funding success is statistically insignificant, implying offline authentication is able to reduce taste-based discrimination[18]. Besides, the additional offline process provides a good opportunity to trustworthy first-time borrowers. Fuster et al. (2019) suggest first-time mortgage borrowers prefer to have a face-to-face communication with loan officers rather than directly applying online as they are less familiar with the process involved. After talking with local loan officers, the loan officers not only will instruct borrowers with application process, but warn them the consequences of not paying the loan back as well. For example, in the case of default, the platform may first call the borrower and then a collection agent will step in (Tao et al., 2017). So, in order to avoid such situation, the borrower will try their best to repay the loan. Through face-to-face interaction with the

[18] Becker (1957) introduced taste-based discrimination in which a principal simply has a preference for working with one type over the other.

borrowers, loan officers can also verify soft information such as good characters, which makes higher possibility of repayment.

The financial excluded might benefit from additional offline authentication as well. As mentioned above, low funding success rate of the financial excluded is due to high information asymmetry so lenders are not willing to provide credit. The offline authentication mitigates asymmetric information significantly. Therefore, the financial excluded with offline authentication can be given a higher credit score, thereby increasing the likelihood of funding success. If the financial excluded has not conducted offline authentication, they might still be given a low credit score and have less chance to succeed in funding campaign. Taken together, the offline authentication moderates the financial excluded and credit score such that this relationship is positive when the financial excluded has conducted offline authentication but negative when they has not. In the end, the high (low) credit score will lead to high (low) funding success rate. Against this background, we therefore hypothesize:

H2: *offline authentication attenuates the mediated effect of credit score in the relationship between the financial excluded and funding success*

Uploading education information is an alternative way for P2P borrowers to mitigate information asymmetric. The highly educated borrowers are expected to have stable employment and higher income and they are perceived less risky borrowers (Xu et al., 2018). Therefore, the education attainment can serve as an effective signal to select a more creditworthy borrower. More importantly, it is costly and difficult to manipulate (Xu et al., 2018). The evidences from P2P lending market in China generally support this view. Using data from Renrendai, Chen and Ning (2013) and Xu et al (2018) document that borrowers who have obtained a higher education degree have higher likelihood to have their loan request funded. Huang et al (2020) uncover borrowers with higher education qualification are associated with lower likelihood of default, fewer overdue payments and smaller overdue amount based on data from Xinxindai. Moreover, Li and Hu (2019) suggest not only levels of education but the institutions matter in China as well. Based on data from Renrendai, they find borrowers who have degree from top ranking universities are less likely to default and their results are robust using instrumental variable method.

In addition, educated people are more like to make prudent financial decisions (Fernandes et al., 2014). If the interest rate is too high, educated borrowers are less likely to take loan offered by these predatory lenders (Xu et al., 2018). Using a meta analysis, Fernandes et al (2014) find financial education leads to positive savings. In contrast to the prudent financial decision, the ill-advised financial decision by borrowers are the one of the antecedents of delinquency of mortgage loans and the effect is pronounced among less creditworthy borrowers (Agarwal et al., 2010). More than 140 studies have indicated that education is associated with better financial outcomes (Miller et al., 2014). For example, Cole et al (2016) suggest High school mathematics training is beneficial to students, which enables greater levels of financial participation, more investment income and better debt management. For the financially disadvantaged cohort, debt management skills from the education alleviates a range of adverse debt behaviors (French and McKillop, 2016). These skills eventually become a personal finance-specific form of human capital (Huston, 2015).

Given the benefit from the better education for financially disadvantaged borrowers mentioned above, we predict education qualification plays a vital role in the association between the financial excluded and credit score. As the financial excluded with higher education qualification are prudent, they must be ready to commit further financial responsibility and know how to manage their debt before the loan application. The platform therefore will value educated borrowers by awarding them a higher credit score. Taken together, the educational attainment moderates the financial excluded and credit score such that this relationship is positive when the financial excluded has higher education qualification but negative when they has not. The higher score then translates to higher likelihood of funding success. Against this background, we therefore hypothesize:

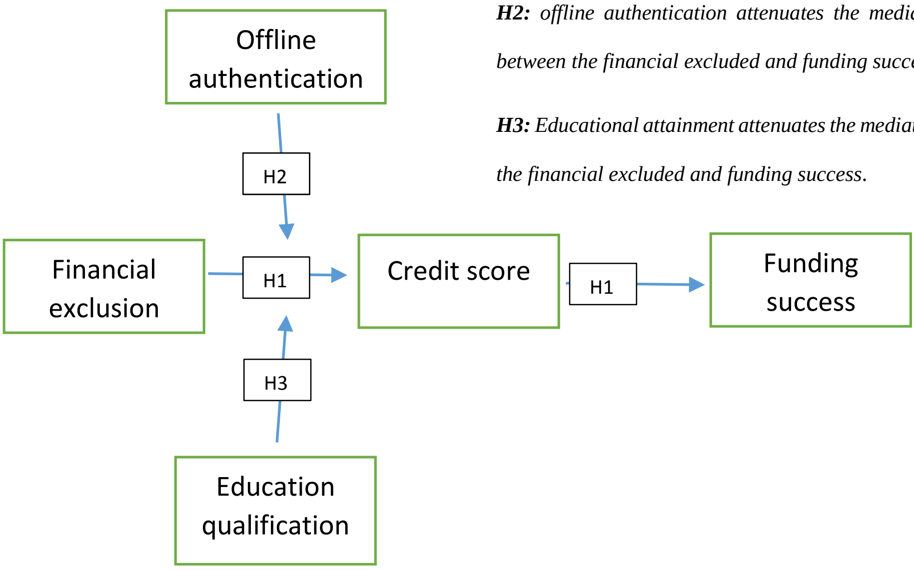
H3: *Educational attainment attenuates the mediated effect of credit score in the relationship between the financial excluded and funding success.*

Figure 2. Conceptual framework

H1: *The credit score mediates the relationship between the financial excluded and funding success.*

H2: *offline authentication attenuates the mediated effect of credit score in the relationship between the financial excluded and funding success.*

H3: *Educational attainment attenuates the mediated effect of credit score in the relationship between the financial excluded and funding success.*



4.3 Research Setting

In 2013, the Chinese government announced the "inclusive finance" development strategy, which is to make financial services accessible to Small and Medium Enterprises (SMEs) and rural population that have traditionally been denied credit by banks. According to the government white paper, the development strategy encourages financial innovations to produce financial products. The ultimate goal is to build a well-established financial system that can meet the needs of a larger group of people. Now the financial system is regarded as the privilege of the State-Owned Enterprises (Fungacova and Weill, 2014). Fintech innovation in P2P credit market may be able to response the call for "inclusive finance" in China by offering great opportunities to those who are financially excluded as the market is consist of a larger number of credit providers and excluded borrowers who can participate freely in this market. Policy makers give priority to financial exclusion as greater access to financial services is beneficial to economic activities and population's welfare. However, financial accessibility remains an unresolved issue in China (Chen and Jin, 2017). In developed economies, people obtain credit according to their personal credit risk and readily available information from commercial banks or government lending agencies (Chai et al., 2019). Yet, in developing economies borrowers' information is limited and accessing borrowers' creditworthiness is difficult (Grant, 2007). Credit information such as credit scores is also not sufficient in China. Only 350 million citizens have credit histories, less than one-third of the adult population (Economist, 2016). In the US, the credit scores have awarded to 89% of adults (Economist, 2016). Without such information, individuals find difficult to get access to credit. Moreover, with the reform of state-owned banks and rural credit cooperatives in China, tens of thousands of branches and entities closed, which leads to more difficulties in lending and borrowing (Sparreboom and Duflos., 2012). Normally to obtain credit from formal institutions, collateral assets and steady employment are required. Yet, the conditions are difficult to meet. This is especially true for those people with low pay job and low levels of education (Beck et al, 2006). Those also are the people who need credit most. Financial institutions are yet not willing to provide loans to low income people even though they have met those criteria because the loan they required is often too small to be profitable (Johnston and Morduch, 2008). Fungacova and Weill (2014) suggest individuals in China find difficult to get access to formal credit as formal financial institutions target state-owned enterprises. There is therefore a credit gap for disadvantaged borrowers especially for very small loans (Mills and McCarthy, 2016) As a result, these difficulties make the vulnerable such as rural and low-income households financially excluded.

The inception of P2P lending tries to fills the funding gap by offering an alternative access to financial resources. First, most of P2P lending platforms do not require collaterals and the lender cannot ask for collaterals. A number of borrowers without collateral benefit from the natural of uncollateralized lending. Second, ideally, P2P transactions are completed online. Borrowers submit their documentations to the website and lenders will review and evaluate the information by themselves to make the decision. The pure online process is not only beneficial to the borrowers from rural areas but to the lenders who want to invest remotely.

4.4 Empirical design

In this section, we first test the mediating effect of credit score on the association between financial exclusion and funding success in the presence of unique offline authentication business model. Following regressions are applied to test our proposed hypotheses.

$$Credit\ Score = \alpha_0 + \alpha\ Exclusion_i + \ \Psi_3 Borrower\_info_i + \varepsilon_i \quad (1)$$

$$Prob(Success_i = 1) = \beta_0 + \beta_1\ Credit\ Score_i + \ \Gamma_4 Borrower\_info_i + \Phi_2 Loan\_info_i + \mu_i (2)$$

Where i indicates ith listings. The data is from Renrendai, one of the biggest P2P platforms in China, between 1 Jan 2015 and 31 Dec 2015. There 400895 listings over our sample period. After deleting missing values, we, finally, get in total 264367 listings. Among them, 113381 listings are fully funded. Each listing has a set of information available to potential investors when they made their investment decisions. These information include (1) the loan characteristics such as interest rate, credit score, loan amount, loan duration, and loan description, (2) borrower's information such as gender, age, education qualification, marital status, personal income and borrowing history and so on.

The dependent variable, *Funded*, in Eq(2) is a dummy variable, which takes 1 if the listing is funded successfully and 0 otherwise.

The key independent variable is *Exclusion* which takes 1 if the borrower claims their income is below 5000RMB and didn't get mortgage or car loans before. The low-income households have long been considered financial exclusion (Ibtissem and Bouri, 2013). The average income of our sample is around 5500RMB so we set our threshold below the average. If the borrower has obtained the mortgage or car loan before, they therefore are able to get access to financial services and are not excluded. Hence, only the borrowers who have a low income and don't have mortgage or car loan are regarded as financial exclusion in this study.

The mediator, *credit score*, is a continuous variable, which ranges from -39 to 223. After borrowers upload their information to Renrendai, Renrendai will award them a credit score based on borrowers' creditworthiness. To make credit score understandable to all the investors, Renrendai created their own credit rating system. Each credit score has a corresponding credit rating. Their one-to-one relation is as follows: AA: above 210; A: 180-209; B: 150-179; C:130-149; D: 110-129; E: 100-109; HR: below 99. We choose credit score not credit rating because we would like to disentangle how the platform not investors treats financially excluded borrowers and the platform uses credit score to determine their creditworthiness. Moreover, credit score is more specific and detailed than credit rating which enables us to capture a small variation in the borrower's creditworthiness.

Offline and *education* are moderators. *Offline* takes 1 if the borrow has conducted offline authentication and 0 otherwise. *Education*, a categorical variable indicating the education level from 0 (high school or lower) to 3 (master or above).We then include two interaction terms *offexclu*, product of *offline* and *exclusion*, and *eduexclu*, product of *education* and *exclusion* to

Eq(1) separately. We would like to test the extent to which offline authentication and borrowers' education qualification can increase/reduce the credit score of financially excluded borrowers, and as a result, increase/reduce funding probability.

We follow the P2P literature to control for a number of loan and borrow level variables that have an influence on funding success (Liu et al, 2015; Lin et al, 2013; Freedman and Jin, 2017; Dorfleitner et al, 2016; Li and Hu, 2019; Ding et al, 2018). *Loan_info* is a set of loan level control variables it includes (1) *Interest*, the interest rate of the listing ranging from 7% to 13.2%; (2) *Months*, The number of months the borrowers would like to pay back the loan ranging from 3 to 48 months; (3) *Lnamount*, the natural logarithm of the loan amount requested by borrowers and (4) *Wordcount*, the number of words written by the borrower in description. *Borrower_info* is a set of borrower level control variables. *Borrower_info* includes (1) *Application*, a dummy variable which takes 1 if the borrower has applied loan in the platform before and 0 otherwise; (2) *Entrepren* a dummy variable which takes 1 if the loan is used to serve business activity and 0 otherwise; (3) *Age*, the borrower's age in years; (4) *Male*, a dummy variable which takes 1 if it is male borrower or 0 otherwise; (5) *Married*, a dummy variable which takes 1 if the borrower is married and 0 otherwise.

This study adopts structural equation modeling (SEM) to test the mediation effect of credit score in the relationship between the financial excluded and funding success. There are many studies that recommend SEM approach to test mediation effect (Iacobucci et al., 2007; Zhao et al., 2010; Cho and Pucik., 2005). This method also enables us to check the direct negative effect of vulnerable borrowers on P2P funding success (Komarova Loureiro and Gonzalez., 2015). Using SEM instead of three separate regressions proposed by Baron and Kenny (1986) can better control for measurement errors which might result in under- or over-estimation of mediation effects (Shaver, 2005). However, the nature of our dependent variable (funded is a dummy variable) makes a linear SEM ill-suited. Applying a linear model with dummy dependent variable will lead to biased results. Hence, we use Generalized SEM (GSEM) model that allows binary outcome to fit our proposed regressions (Kaplan and Vakili., 2014). We apply linear regression to first part of analysis (determinants of credit score) and logit regression with a dummy dependent variable to second part (determinants of funding success). Bootstrapping is used to estimate standard error and confidence intervals for the indirect effects.

4.5 Results

Table 1 shows the descriptive statistics of the variables used. To test H1, We follow the mediation testing procedure introduced by Zhao et al (2010). They suggest to test the indirect effect $\alpha * \beta$ (the coefficient of Exclusion in Eq(1) multiplies the coefficient of credit score in Eq(2) in our case) using a Bootstrap method. The result is shown in Table 2 Panel B. With a 1000 repetitions Bootstrap method, coefficient of Exclusion on Funding success through credit score is -0.0053 and it is significant at 1% level (P<0.001). The results support the mediating role of credit score between financial exclusion and funding success so H1 is supported.

Table 165 Descriptive Statistics

Variable	N	Mean	S.D	Min	Max
CreditScore	264367	84.78	80.03	-39	223
Exclusion	264367	0.280	0.450	0	1
Offline	264367	0.400	0.490	0	1
Education	264367	0.950	0.770	0	3
Age	264367	33.20	7.940	20	63
Male	264367	0.800	0.400	0	1
Married	264367	0.590	0.490	0	1

application	264367	0.230	0.420	0	1
Entrepren	264367	0.100	0.300	0	1
funded	264367	0.430	0.490	0	1
interestrate	264367	0.120	0.010	0.070	0.130
months	264367	22.27	10.21	3	48
Lnamount	264367	10.65	0.980	6.910	13.12
wordcount	264367	45.63	31.74	0	363

Table 176. Correlation Coefficient

	CreditScore	exclusion	offline	education	income	age	male	married
CreditScore	1							
exclusion	-0.222***	1						
offline	0.965***	-0.214***	1					
education	0.114***	-0.082***	0.083***	1				
income	0.173***	-0.627***	0.165***	0.052***	1			
age	0.401***	-0.194***	0.404***	0.013***	0.217***	1		
male	-0.178***	0.017***	-0.184***	-0.034***	-0.017***	-0.084***	1	
married	0.230***	-0.171***	0.227***	-0.033***	0.162***	0.347***	-0.047***	1
application	-0.380***	0.096***	-0.444***	-0.026***	-0.069***	-0.200***	0.105***	-0.100***
funded	0.951***	-0.211***	0.928***	0.105***	0.164***	0.389***	-0.171***	0.223***
interestrate	-0.357***	0.032***	-0.355***	-0.025***	-0.148***	-0.124***	0.029***	-0.065***
months	0.599***	-0.170***	0.623***	0.107***	0.011***	0.278***	-0.155***	0.159***
Lnamount	0.308***	-0.373***	0.323***	0.143***	0.393***	0.305***	-0.119***	0.209***
wordcount	0.569***	-0.181***	0.583***	0.078***	0.204***	0.280***	-0.108***	0.157***

Table 187. Mediation Testing

Panel A Structural Equation Modeling

DV	Funded	VIF	CreditScore	VIF
exclusion	-0.684***	1.21	-2.098***	1.08
	(-22.71)		(-22.13)	
offline	0.633***	19.53	160.1***	1.52
	(6.46)		(1488.44)	
education	0.113***	1.06	3.444***	1.02
	(6.81)		(58.89)	
age	0.0403***	1.35	0.115***	1.31
	(17.52)		(22.93)	
male	-0.0636	1.04	-0.214*	1.04
	(-1.71)		(-2.41)	
married	0.141***	1.17	1.455***	1.16
	(4.82)		(16.39)	
application	-1.310***	1.31	11.51***	1.25
	(-40.94)		(80.60)	
Entrepreneurship	0.288***	1.04	0.233	1.01
	(6.50)		(1.57)	

interestrate	-13.60***	3.23		
	(-5.06)			
months	0.0308***	4.92		
	(8.20)			
Lnamount	-1.103***	1.61		
	(-70.35)			
CreditScore	0.0554***	15.52		
	(105.06)			
constant	6.112***	Mean VIF 4.20	11.08***	Mean VIF 1.17
	(20.47)		(56.82)	
N	264367		264367	

Panel B Mediation Testing with Bootstrapping (1000 repetitions)

Mediator	Coefficient (a x b)
CreditScore	-0.00533***
	(-3.19)

Panel C Indirect, Direct and Total Effects

	log odds ratio	percentage
indirect	-0.995	0.976
direct	-0.024	0.024

total	-1.019	1

* $p<0.05$; ** $p<0.01$; *** $p<0.001$; Z statistics are in parentheses

Although the mediation effect indeed exists, if the indirect effect of exclusion relative to total effect is too small, to investigate indirect effect is therefore unnecessary. To explore the indirect and total effect, we follow methods suggested by Erikson et al (2005), Buis (2010) and Berry Jaeker et al (2020) to calculate the proportion of indirect effect in a logit outcome equation (Stata command: *ldecomp*). The results are shown Table 2 Panel C. The first column is log odds ratio and the second shows the percentage. The indirect effect is 97.6% of the total effect.

To test H2 and H3, we follow Preacher et al (2007)'s method to test a moderated mediation. When the offline authentication moderates the association between financial exclusion and credit score, with a 1000 repetitions Bootstrap method, the coefficient of conditional indirect effect of financial exclusion on funding success is from -0.0609 (without offline authentication, $P<0.000$) to 0.0078 (with offline authentication, $P<0.011$). This means offline authentication model can offset the negative impact of financial exclusion on funding success so H2 is supported.

We then test moderated mediation effect of education qualification and the results are presented in Table 3. With a 1000 repetitions Bootstrap method, the negative coefficient of conditional indirect effect of financial exclusion on funding success is gradually decreasing along with an increase in education qualification. When the borrowers hold a high school or lower certificate (education = 0), junior college certificate (education= 1) and bachelor degree (education= 2), the coefficients are -0.0642 ($P<0.000$), -0.0387 ($P<0.000$) and -0.0131 ($P<0.012$) respectively. The coefficient is 0.0123 but it is not statistically significant ($P< 0.128$), if the borrowers' have a master or doctorate degree. In sum, the negative effect of financial exclusion on funding success will attenuate when the borrowers hold a higher education qualification so H3 is supported

Table 198. Moderated Mediation with Bootstrapping (1000 repetitions)

Moderator	Level	Coefficients	SE	Z	P
offline	0	-0.0609	.0037308	-16.33	0.000
	1	0.0078	.0030863	2.54	0.011
Education	0	-0.0642	.0037004	-17.37	0.000
	1	-0.0387	.0031305	-12.37	0.000
	2	-0.0131	.005219	-2.52	0.012
	3	0.0123	.0081245	1.52	0.128

4.6 Robustness and Further analysis

We use the number of bids as alternative dependent variable to check the robustness of our results. As the number of bids is a continuous variable, we can use SEM for testing (Stata command: *sem*). The results are shown in Table 4.

Table 19. Robustness Check

Panel A Structural Equation Modeling		
DV	Funded	CreditScore
exclusion	-2.052***	-2.098***
	(-9.36)	(-22.93)

offline	8.277***	160.1***
	(10.18)	(1599.83)
education	1.397***	3.444***
	(11.64)	(66.48)
age	0.157***	0.115***
	(11.91)	(20.00)
male	-3.004***	-0.214*
	(-13.17)	(-2.14)
married	-0.717***	1.455***
	(-3.62)	(16.70)
application	-0.861***	11.51***
	(-3.53)	(109.57)
Entrepreneurship	-0.388	0.233
	(-1.25)	(1.72)
interestrate	-2336.1***	
	(-141.90)	
months	2.278***	
	(116.47)	
Lnamount	6.652***	

	(56.82)	
wordcount	-0.0860***	
	(-23.99)	
CreditScore	0.0623***	
	(14.06)	
Constant	179.2***	11.08***
	(84.96)	(51.39)
N	264367	264367

Panel B Mediation Testing with Bootstrapping (1000 repetitions)	
Mediator	Coefficient (a x b)
CreditScore	-0.1308***
	(-11.99*)*

* p<0.05; ** p<0.01; ***p<0.001; Z statistics are in parentheses

Offline authentication is a new initiative service provided by the platform to alleviate severe asymmetric information in China. Although it has been criticized that authenticated loans are not P2P loans, it is still welcomed by investors and investors are more willing to make a contribution to such loans. From excluded borrowers' perspective, we show that offline authentication positively moderates the linkage between financial exclusion and credit score and therefore increase the funding probability. Given the positive effect of offline authentication on funding outcome, it is necessary to disentangle the factors that influence borrowers' decision to conduct offline authentication. The following model is used.

$$\text{Prob}(Offline_i = 1) = \beta_0 + \beta_1 Branch_i + \beta_2 Exclusion_i + \Gamma_2 Borrower\_info_i + \varepsilon_i$$

Table 5 shows the results of Eq (3). Model (1) and (2) are the probit and logit estimation respectively. The coefficient of *branch* is 1.755 and it is statistically significant at the 1% level. The coefficient of logit model is 3.183 and it is again statistically significant at the 1% level. If there is a local branch in borrowers' city, they are more likely to conduct offline authentication.. The result implies cities with better financial infrastructure can give P2P borrowers a better opportunity to obtain credit. The coefficients of *Exclusion* are -0.310 and -0.497 in probit and logit model respectively. Both are significant at the 1% level, suggesting that the financial excluded borrowers are less likely to conduct offline authentication. The reason probably is the borrowers who are financially excluded are reluctant to have face-to-face interaction with other people, which is consistent with the finding that financial exclusion may lead to social exclusion (Wilson, 2012). Psychological barriers and the feelings among the poor that they are not qualified to financial services might result in financial exclusion (Solo, 2008). *Entrepren* and *Application* are both negative and significant at the 1% level in both estimation. The borrowers who use their loan to fund their business or have applied the P2P loan before are less likely to undergo offline check than other borrowers. Other control variables such as *education, age,* and *married* have a positive impact on borrowers' offline authentication choice. These imply the borrowers who have better education qualification, older age and been married are more willing to undergo offline check before listing online. Compared to female borrower, male seems more risk taking so they are more likely to take a chance and directly apply online. High income borrowers seem to be confident with their qualification and are more willing to pure online application.

Table 200. Determinants of borrowers' willingness to offline authentication

	(1)	(2)
	Offline	Offline
Lnamount	0.311***	0.567***
	(0.005)	(0.008)
exclusion	-0.310***	-0.497***
	(0.010)	(0.016)
branch	1.755***	3.183***
	(0.012)	(0.023)
education	0.057***	0.093***
	(0.004)	(0.008)
income	-0.067***	-0.111***
	(0.004)	(0.006)
age	0.055***	0.097***
	(0.000)	(0.001)
male	-0.423***	-0.724***
	(0.008)	(0.014)
married	0.311***	0.521***
	(0.007)	(0.012)
application	-2.752***	-5.651***
	(0.034)	(0.075)
Entrepreneurship	-0.514***	-0.928***
	(0.012)	(0.021)
_cons	-6.225***	-11.299***
	(0.054)	(0.098)
N	264367	264367
pseudo R-sq	0.460	0.464

* p<0.05; ** p<0.01; ***p<0.001; Robust Standard error are in parentheses

4.6 Discussion

4.6.1 theoretical contribution

Although researchers are interested in the association between fintech and financial exclusion, the association remains unclear. This study tries to answer this question by investigating the funding performance of the financial excluded borrowers in a large P2P lending platform. Using loan-level data, we find credit score negatively mediates the financial excluded and funding success. We also find offline authentication and borrowers' education background positively moderate the linkage between the financial excluded and credit score. Finally, we document the financial excluded are less likely to conduct offline authentication even though offline authentication give them a better opportunity to obtain the loan.

We contribute to extant literature in following aspects. First, current P2P studies focus on direct effect on funding outcomes but ignore the indirect effect and the underlying channel (Freedman and Jin, 2017; Dorfleitner et al, 2016; Herzenstein et al, 2011; Michels, 2012). The reason is probably because in the US or other major P2P market, the P2P platform plays a little or no role in the investors' decision making process. However, in China, due to lack of nationwide-recognized credit rating system, investors rely heavily on the credit score provided by the P2P platform. Moreover, the funding decision is made very quickly. 25% of campaigns are fully funded within 42 seconds and 75% of them get funded in less than 180 seconds (Liao et al., 2019). If investors take a longer time to make the decision, investment opportunity is gone. Due to time constraint, investors tend to only attend a few salient information. Credit rating which is given by the platform therefore plays a vital role in the relationship between funding success and the financial excluded. We find credit score in fact mediates the association between the financial excluded and funding campaign outcome. Given the important role of P2P lending in financial exclusion and inconclusive findings offered by existing literature (Komarova Loureiro and Gonzalez., 2015; Lin and Viswanathan, 2015), we contribute to the

P2P literature by providing evidence that there is a negative mediation effect of credit score for the financial excluded on funding success.

Second, we contribute financial exclusion literature. The discussion about alleviating financial exclusion has a long history. Whether Fintech is able to combat financial exclusion remains unclear. We investigate the extent to which the financial excluded can benefit from one of the important fintechs, P2P lending. We find although the financial excluded borrowers are less likely to be funded, offline authentication, a new business model invented by Chinese P2P firms, and their education qualification can positively moderate the linkage between the financial excluded and credit score and therefore increase probability of being funded. Our results also suggest although offline authentication can provide additional benefit to the excluded borrowers, they remain reluctant to conduct authentication. The reason probably is the borrowers who are financially excluded are reluctant to have face-to-face interaction with other people, which is consistent with the finding that financial exclusion may lead to social exclusion (Wilson, 2012). Lastly, the financial infrastructure such as local branch plays a positive role in encouraging the borrowers to do the offline authentication. The uneven distribution of financial infrastructure makes the financial exclusion even worse in the areas without financial infrastructure.

4.6.2 Implications

The study 3 offers several practical implications. The study suggests that the indirect effect of the financial excluded through credit score plays a major role in funding performance instead of direct effect. It is therefore possible to allay financial exclusion by improving the credit score system. For example, in recent years, big data offers opportunity to incorporate digital footprints such as the website that the borrowers visited and the mobile they used (i.e. IOS or Android) in assessing borrowers' creditworthiness. Such big data provides more comprehensive and accurate credit information than traditional data (Berg et al., 2020). It also may give the financial excluded a better chance when the platforms use it to determine their credit scores. Furthermore, the results from study 3 show offline authentication and human capital such as education attainment can improve the likelihood of funding success significantly and the establishment of offline branch can lead to higher probability of offline authentication. The offline branch is hence considered an important financial infrastructure that can promote financial inclusion. Given the benefit provided by offline authentication, it is better for platforms to adopt a hybrid operating model that combines both online and offline business, as offline business will give the financial excluded a better opportunity. In addition, to mitigate financial exclusion, the government can subsidize platforms with offline business due to higher cost of the business model. Given the inception of inclusive finance development strategy, subsidies to these P2P platform are also the best use of the funding. Lastly, the education is beneficial to the financial excluded, which sends a good signal to lenders. The platforms can therefore offer some online courses to borrowers so they will have a better understanding of their future financial commitment provided that borrowers often rush to take a loan decision without thinking their ability of paying back.

4.7 Conclusion

This study examines the association between financial exclusion and P2P funding success. We find there is a negative indirect effect of financial exclusion and P2P funding success through borrowers' credit score. Moreover, the indirect effect takes account for about 98% of total effect, which is consistent with our prediction that credit score is an important channel between financial exclusion and funding success. The result supports the argument that fintech innovation doesn't help mitigating financial exclusion. We then explore the moderating role of offline authentication and education attainment in the relationship between financial exclusion and funding success through credit score. The results suggest the negative indirect effect of financial exclusion on funding success through credit score is overturned if excluded borrowers have conducted offline authentication or have a higher education attainment (master or above). Lastly, we investigate the determinants of borrowers' willingness to conduct offline authentication. Excluded borrowers are less likely to conduct offline authentication though their application will benefit from offline authentication. In addition, local branch plays a significant role in offline authentication decision.

Chapter 5. Conclusions

This thesis investigates in different aspects of P2P lending. First, drawing on signaling and message framing theory, the finding suggests that positively-framed message as a cheap signal can be effective in increasing the probability of a loan being funded. We also find overly-positive message has a negative impact on funding success. Moreover, the finding shows cheap signal (message framing) and costly signal (e.g., credit rating) complement to each other in determining the likelihood of funding success.

Second, lenders' decision can be influenced by psychological perspective. We connect P2P literature and psycholinguistic literature and explain how psychological distancing measured by linguistic style influences lenders' decision on funding campaign. The finding shows psychological distancing is inversely related to P2P funding success. In addition, language intensity negatively strengthens the negative relationship between psychological distancing and funding success. Our findings are in line with the notion that psychological distancing is associated with negative personal outcomes (Simmons et al, 2005; Revenstorf et al, 1984).

Third, there is a long debate of effectiveness of fintech on financial exclusion given that the emerging of fintech is to make more people financially inclusive. We find a negative indirect effect of financial exclusion on funding success via credit score. Moreover, the findings also suggest offline authentication and education attainment positively moderates the association between the financial excluded and credit score and thus negative indirect effect of financial exclusion on funding success is mitigated. Given the benefit of the offline authentication, we examine the determinants of offline authentication. The finding suggests borrower in a city with better financial infrastructure are more likely to conduct offline authentication. However, financial excluded borrowers are less likely to conduct offline authentication.

Three empirical chapters shed new light and provide insights on different aspects of P2P market. The conclusion chapter is to integrate and summarize the key contributions and implications of these chapters. Finally, it offers potential avenues for further studies.

5.1 key contributions

This thesis makes several main contributions to the literature. First, it contributes to the role of language in entrepreneurial finance. The study 1 and study 2 discuss both positive and negative language and suggest the language plays an important role in funding campaign. Second, it makes contributions to the debate of financial innovation and financial inclusion. The results show the financial innovation not necessarily improves financial inclusion. Following paragraphs will elaborate the contributions in details.

The findings of study 1 (chapter 2) show positively framed messages related to trust have a positive influence on funding performance but not on number of bids. The study reconciles mixed findings in the literature regarding online reviews (e.g. Ludwig et al., 2013; Salehan and Kim, 2016) by suggesting that message framing matters in context of online environment. Drawing upon message framing and costless signaling theory, the study shows language can be an effective signal even though it does not cost too much. The findings are also consistent with marketing literature which posits framings have a larger effect when consumers have limited knowledge about the product (Chang, 2007). In P2P platforms, lenders are barely familiar with the borrowers or the projects that they are going to invest. Moreover, the study contributes to traditional signaling theory which suggests cheap signals are not able to create a separating equilibrium between better and worse signalers (Balvers et al., 2014). The difficulty for lenders to assess costly signals may justify the importance of language as a cheap signal in P2P lending decisions. Importantly, the findings show cheap signals complement costly signals in achieving funding success. This implies costly signals and costless signals are complemented to assess the quality of the potential borrowers, which is consistent with a small number of studies that show costly signals are beneficial on funding success while cheap signals enhance the benefit by promoting communications between entrepreneurs and funders (Davis and Allison, 2013). Such communication is particularly important in virtual environments, in which impression formation based on face-to-face interactions and experiences is lacking. The finding responses the call for research on interaction of various signals on outcomes (Anglin et al., 2018). The extant studies so far tend to examine external signals in isolations (Anglin et al., 2018). The costly signals mitigate the risks of cheating and any costs associated with misleading signals (Connely et al., 2011). However, the costs of relying exclusively on costly signals and assessing them may be high. As such, the complementary use of cheap signals may be cost-effective for senders of such signals.

The second study bridges the psycholinguistics literature and P2P lending literature and finds psychological distancing negatively affects P2P funding performance. The finding is related to psycholinguistics literature that posits psychological distancing has a negative impact on interpersonal outcome (Simmons et al, 2005; Revenstorf et al, 1984). Moreover, it shows the language intensity strengthens the negative relationship between funding success and psychological distancing, which is also consistent with the argument that suggest showing an extreme position is negatively related to perceived source competence and reduces persuasive power (Buller et al., 1998). The study 2 contributes the literature on following aspects. First the paper is closely associated with the role of language in crowdfunding (Parhankangas and Renko, 2017; Anglin et al., 2018; Herzenstein et al., 2011; Allison et al., 2013; Majumdar and Bose, 2018).

The intent of the literature is to show how borrowers improve their funding performance by strategically using language. Drawing on message framing or signaling theory, studies focus on narratives that have been written on purpose (Huang et al., 2020; Anglin et al., 2018; Block et al., 2018). They ignore the fact that borrowers sometime write something unconsciously but this can also affect lenders' perception. Drawing on psycholinguistics literature, study 2 shows unconscious use of certain language styles resulting in high psychological distancing has a negative effect on funding performance. Second, the study contributes to language expectancy theory. Existing literature adopts language expectancy theory to test how message features such as opinionated language, sequential message, language intensity, message contents, lexical complexity or fear arousing appeals to meet the receivers' expectation (Averbeck and Miller, 2014; Foste and Botero, 2012; Jensen et al., 2013). They tend to neglect relatively invisible style words. The study 2 fills the research gap by suggesting that receivers have expectation on linguistic styles as well. The linguistic styles that show similarity and clarity of speech are more likely to meet the expectation of the borrowers. Lastly, existing studies examine linguistic styles and language intensity separately. However, language intensity that is closely related to tone of communication cannot be separately from linguistic styles. Recipients form their opinion about quality of messages from both linguistic styles and tone of communication simultaneously. Therefore, investigating the interaction effect of language intensity and linguistic style is necessary. the results suggest that language intensity negatively strengthen the negative association between the number of second person pronouns and funding success but has no significant effect on the relationship between negations and funding success. Crowdfunding is an interdisciplinary area, which requires not only purely finance perspectives but strategy and management, psychology and sociology as well. Cumming and Johan (2017) call for the synthesis of various research areas to advance theories and empirical testing in crowdfunding. In this end, we response their call by integrating psycholinguistics literature and P2P study.

The study 3 contributes to two streams of literature. First, it makes contributions to P2P lending literature. Existing P2P studies focus on direct effect on funding performance but overlook indirect effect and underlying channels (Freedman and Jin, 2017; Dorfleitner et al, 2016; Herzenstein et al, 2011; Michels, 2012). The reason is probably because in the US or other major P2P market, the P2P platform plays a little or no role in the investors' decision making process. However, in China, due to lack of nationwide-recognized credit rating system, investors rely heavily on the credit score provided by the P2P platform. Moreover, the funding decision is made very quickly. 25% of campaigns are fully funded within 42 seconds and 75% of them get funded in less than 180 seconds (Liao et al., 2019). If investors take a longer time to make the decision, investment opportunity is gone. Due to time constraint, investors tend to only attend a few salient information. Credit rating which is given by the platform therefore plays a vital role in the relationship between funding success and the financial excluded. We find credit score in fact mediates the association between the financial excluded and funding campaign outcome. We contribute to the P2P literature by providing evidence that there is a negative mediation effect of credit score for the financial excluded on funding success. Second, the study contributes financial exclusion literature. The discussion about alleviating financial exclusion has a long history. Whether Fintech is able to combat financial exclusion remains unclear. We investigate the extent to which the financial excluded can benefit from one of the important fintechs, P2P lending. We find although the financial excluded borrowers are less likely to be funded, offline authentication, a new business model invented by Chinese P2P firms, and their education qualification can positively moderate the linkage between the financial excluded and credit score and therefore increase probability of being funded. Our results also suggest although offline authentication can provide additional benefit to the excluded borrowers, they remain reluctant to conduct authentication. The reason probably is the borrowers who are financially excluded are reluctant to have face-to-face interaction with other people, which is consistent with the finding that financial exclusion may lead to social exclusion (Wilson, 2012). Lastly, the financial infrastructure such as local branch plays a positive role in encouraging the borrowers to do the offline authentication.

5.2 Implications

The study 1 suggests that message framing as a cheap signal can be effective in enhancing the likelihood of a project being funded. In particular, messages framed in a way to signal trustworthiness of potential borrowers have greater likelihood for funding success. Potential borrowers signaling messages for P2P lenders need to frame their messages with words such as honesty, integrity, credence, and reliable which signal their trustworthiness quality. Besides trustworthiness, potential borrowers can also display other qualities through language to attract investors and increase the potential for funding success. For example, signaling words associated with the agility, proactiveness, ambiguousness, empathy and network size of a potential borrower may support lenders' decisions on whether the borrower is worth investing. In this sense, message framing can be very relevant. Furthermore, as suggested by our findings, to enhance the odds of funding success, costly signals such as information on the credit rating of a potential borrower can be complemented by suitably framed message embedded in the description of the project. This is because, when a borrower communicates their credit rating through objectively assessed evidence, their continuous emphasis on their credit history may only have a supplementary role, and thus have marginal additional influence on their funding success. If borrowers seeking for funds can enhance their communication mode through the use of positive messages, their qualities, which can be objectively assessed, would be supported through language, and therefore increase the likelihood of being funded

The study 2 provides some practical implications. For borrowers, it is important to avoid psychological distancing with lenders through linguistic styles and tone of communication. It is likely that borrowers use certain linguistic styles unconsciously which distances them from the crowd. However, this can be avoided if the borrowers take a longer time to prepare their loan application. For example, borrowers should reduce the use of second person pronouns and negations. Especially in China, the use of second person pronouns, “ni” shows disrespectful. Recipients are confused if high frequency use of negations in the communication because it is difficult to process negations for human beings. In addition, showing an extreme position through intensive language may reduce the probability of funding success further. Lenders may have some doubts about their payback ability in this circumstances. Therefore, less intensive tone is suggested when the borrowers ask for a loan.

The study 3 offers several practical implications. The study suggests that the indirect effect of the financial excluded through credit score plays a major role in funding performance instead of direct effect. It is therefore possible to allay financial exclusion by improving the credit score system. For example, in recent years, big data offers opportunity to incorporate digital footprints such as the website that the borrowers visited and the mobile they used (i.e. IOS or Android) in assessing borrowers’ creditworthiness. Such big data provides more comprehensive and accurate credit information than traditional data (Berg et al., 2020). It also may give the financial excluded a better chance when the platforms use it to determine their credit scores. Furthermore, the results from study 3 show offline authentication and human capital such as education attainment can improve the likelihood of funding success significantly and the establishment of offline branch can lead to higher probability of offline authentication. The offline branch is hence considered an important financial infrastructure that can promote financial inclusion. Given the benefit provided by offline authentication, it is better for platforms to adopt a hybrid operating model that combines both online and offline business, as offline business will give the financial excluded a better opportunity. In addition, to mitigate financial exclusion, the government can subsidize platforms with offline business due to higher cost of the business model. Given the inception of inclusive finance development strategy, subsidies to these P2P platform are also the best use of the funding. Lastly, the education is beneficial to the financial excluded, which sends a good signal to lenders. The platforms can therefore offer some online courses to borrowers so they will have a better understanding of their future financial commitment provided that borrowers often rush to take a loan decision without thinking their ability of paying back.

5.3 Future Directions and Research Agenda

This thesis analyzes the role of language in P2P lending and the association between P2P lending and financial exclusion. Importantly, this study focused on borrower-related issues at the individual level from only the borrower perspective. Future research can use a multi-level study to analyze how individual-related and environmental issues, from both borrower and investor perspective, may affect funding performance of potential borrowers. Another issue is this study focused on one particular website. Future P2P research can try to do some cross-platforms investigations.

Reference

Abe, J., 2011. Changes in Alan Greenspan's Language Use Across the Economic Cycle: A Text Analysis of His Testimonies and Speeches. *Journal of Language and Social Psychology*, 30(2), pp.212-223.

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S. and Evanoff, D., 2010. Learning to Cope: Voluntary Financial Education and Loan Performance during a Housing Crisis. *American Economic Review*, 100(2), pp.495-500.

Aggarwal, S. and Klapper, L., 2013. Designing government policies to expand financial inclusion: Evidence from around the world. *The Journal of Finance*, *56*(3), pp.1029-51.

Ahlers, G. K.C., Cumming, D., Günther, C. and Schweizer, D., 2015. Signaling in Equity Crowdfunding. *Entrepreneurship Theory and Practice*, *39*(4), pp.955-980.

Aktas, N., De Bodt, E., Bollaert, H., & Roll, R., 2016. CEO narcissism and the takeover process: From private initiation to deal completion. *Journal of Financial and Quantitative Analysis*, *51*(1), pp.113-137.

Allison, T. H., McKenny, A. F., & Short, J. C., 2013. The effect of entrepreneurial rhetoric on microlending investment: An examination of the warm-glow effect. *Journal of Business Venturing*, 28(6), pp. 690–707.

Amaeshi, K., 2006. Financial Exclusion, Financial Institutions and Corporate Social Responsibility: A Developing Country Perspective. *SSRN Electronic Journal*,.

Ambrose, B.W., Conklin, J. and Yoshida, J., 2016. Credit rationing, income exaggeration, and adverse selection in the mortgage market. *The Journal of Finance*, *71*(6), pp.2637-2686.

Andres, C., Betzer, A., & Limbach, P., 2014. Underwriter reputation and the quality of certification: Evidence from high-yield bonds. *Journal of Banking & Finance*, *40*, pp. 97-115.

Anglin, A.H., Short, J.C., Drover, W., Stevenson, R.M., McKenny, A.F. and Allison, T.H., 2018. The power of positivity? The influence of positive psychological capital language on crowdfunding performance. *Journal of Business Venturing*, *33*(4), pp.470-492.

Angrist, J. and Krueger, A., 2001. Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives*, 15(4), pp.69-85.

Arya, S., Eckel, C. and Wichman, C., 2013. Anatomy of the credit score. *Journal of Economic Behavior & Organization*, *95*, pp.175-185.

Audi, R., Loughran, T. and McDonald, B., 2016. Trust, but verify: MD&A language and the role of trust in corporate culture. *Journal of Business Ethics*, *139*(3), pp.551-561.

Averbeck, J. M., 2014. Expanding language expectancy theory: The suasory effects of lexical complexity and semantic complexity on effective health message design. *Communication Studies*, 65(1), pp.72–95.

Avey, J. B., Reichard, R. J., Luthans, F., & Mhatre, K. H., 2011. Meta‐analysis of the impact of positive psychological capital on employee attitudes, behaviors, and performance. *Human Resource Development Quarterly*, *22*(2), pp.127-152.

Bachmann, A., Becker, A., Buerckner, D., & Hilker, M. K. M., Lehmann, M. and Tiburtius, P., 2011. Online peer-to-peer lending-a literature review. *Journal of Internet Banking and Commerce*, *16*(2), pp.1-18.

Barakova, I., Bostic, R.W., Calem, P.S. and Wachter, S.M., 2003. Does credit quality matter for homeownership?. *Journal of housing Economics*, *12*(4), pp.318-336.

Baron, R.M. and Kenny, D.A., 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), pp.1173.

Beck, T., Demirgüç-Kunt, A. and Martinez Peria, M., 2008. Banking Services for Everyone? Barriers to Bank Access and Use around the World. *The World Bank Economic Review*, 22(3), pp.397-430.

Becker, G., 2010. *The Economics Of Discrimination*. Chicago: The University of Chicago Press.

BERGER, S.C. and GLEISNER, F., 2009. Emergence of Financial Intermediaries in Electronic Markets: The Case of Online P2P Lending. *Business Research*, **2**(1), pp.39-65.

Bergh, D. D., Connelly, B. L., Ketchen Jr, D. J., & Shannon, L. M. 2014. Signaling theory and equilibrium in strategic management research: An assessment and a research agenda. *Journal of Management Studies*, *51*(8), pp.1334-1360.

Besharat, A. 2010. How co-branding versus brand extensions drive consumers' evaluations of new products: A brand equity approach. *Industrial Marketing Management, 39*, pp.1240–1249.

Beukeboom, C. J., Finkenauer, C., & Wigboldus, D. H. J. 2010. The negation bias: When negations signal stereotypic expectancies. *Journal of Personality and Social Psychology*, 99(6), pp.978–992.

Block, J. H., Hornuf, L., & Moritz, A., 2018. Which Updates During an Equity Crowdfunding Campaign Increase Crowd Participation. *Small Business Economics*, 50(1), pp.3–27.

Boven, L. V., Kane, J., McGraw, A. P., & Dale, J., 2010. Feeling Close: Emotional Intensity Reduces Perceived Psychological Distance. *Journal of Personality and Social Psychology*, 98(6), pp.872–885.

Bowers, J. W., 1963. Language intensity, social introversion, and attitude change. Communication Monographs, 30(4), pp.345–352.

Bowers, John Waite., 1963. Language Intensity, Social Introversion, and Attitude Change. *Speech Monographs*, 30, pp. 345–52.

Boyreau-Debray, G. and Wei, S.J., 2005. *Pitfalls of a state-dominated financial system: The case of China* (No. w11214). National Bureau of Economic Research.

Breuer, W., Knetsch, A. and Salzmann, A., 2018. What Does It Mean When Managers Talk About Trust?. *Journal of Business Ethics*,

Bridges, S. and Disney, R., 2004. Use of credit and arrears on debt among low‐income families in the United Kingdom. *Fiscal Studies*, *25*(1), pp.1-25.

Bruhn, M. and I. Love., 2014. The Real Impact of Improved Access to Finance: Evidence from Mexico. *Journal of Finance*, 69(3), pp.1347-1376.

Bruton, G., Khavul, S., Siegel, D., & Wright, M., 2015. New financial alternatives in seeding entrepreneurship: Microfinance, crowdfunding, and peer‐to‐peer innovations. *Entrepreneurship Theory and Practice*, *39*(1), pp.9-26.

Buller, D. B., Borland, R., & Burgoon, M., 1998. Impact of Behavioral Intention on Effectiveness of Message Features Evidence From the Family Sun Safety Project. *Human Communication Research*, 24(3), pp. 433–453.

Buller, D. B., Burgoon, M., Hall, J. R., Levine, N., Taylor, A. M., Beach, B., … Melcher, C. (2000). Long-Term Effects of Language Intensity in Preventive Messages on Planned Family Solar Protection. *Health Communication*, 12(3), 261–275.

Burgers, C., Beukeboom, C. J., & Sparks, L., 2012. How the doc should (not) talk: When breaking bad news with negations influences patients' immediate responses and medical adherence intentions. *Patient Education and Counseling*, 89(2), pp.267–273.

Burgers, C., Beukeboom, C. J., Sparks, L., & Diepeveen, V., 2015. How (not) to inform patients about drug use: use and effects of negations in Dutch patient information leaflets. *Pharmacoepidemiology and Drug Safety*, 24(2), pp.137–143.

Burgoon, M. Language expectancy theory: Elaboration, explication, and extension. In C.R. Berger and M. Burgoon (ed.), Communication and Social Influence Processes. *East Lansing: Michigan State University Press*, 1995, pp. 29–52.

Burgoon, M., and Miller, G.R. An expectancy interpretation of language and persuasion. In H. Giles and R.N. St. Clair (eds.), Handbook of Language and Social Psychology. *London: Wiley & Sons*, 1985, pp. 51–72.

Burgoon, Michael, Vickie Pauls Denning, and Laura Roberts. 2002. Language Expectancy Theory. In *The Persuasion Handbook: Developments in Theory and Practice*, eds. James P. Dillard and Michael Pfau. Thousand Oaks, CA: Sage, pp.117–36.

Chai, S., Chen, Y., Huang, B. and Ye, D., 2019. Social networks and informal financial inclusion in China. *Asia Pacific Journal of Management*, *36*(2), pp.529-563.

Chang, C-T., 2007. Health-Care Product Advertising: The Influences of Message Framing and Perceived Product Characteristics. *Psychology & Marketing*, *24*(2), pp.143–169.

Chen, D. and Han, C. (2012). A comparative study of online p2p lending in the USA and China. *Journal of Internet Banking and Commerce*, 17(2):1-15.

Chen, J. -Z., and X. Ning., 2013. Empirical Research on the Influence of Personal Information on P2P lending's Success Rate, *Accounting and Finance*, 6, 13–7 (*in Chinese*)

Chen, X., Huang, B. and Ye, D., 2018. The role of punctuation in P2P lending: Evidence from China. *Economic Modelling*, *68*, pp.634-643.

Chen, X., Huang, B. and Ye, D., 2019. Gender Gap in Peer-to-Peer Lending: Evidence from China. *Journal of Banking & Finance*, p.105633.

Chen, X.-P., Yao, X., & Kotha, S., 2009. Entrepreneur Passion And Preparedness In Business Plan Presentations: A Persuasion Analysis Of Venture Capitalists' Funding Decisions. *Academy of Management Journal*, 52(1), pp.199–214.

Chen, Y. and Bellavitis, C., 2020. Blockchain disruption and decentralized finance: The rise of decentralized business models. *Journal of Business Venturing Insights*, 13, pp. 00151.

Chen, Z. and Jin, M., 2017. Financial inclusion in China: use of credit. *Journal of Family and Economic Issues*, *38*(4), pp.528-540.

Cheung, C. M.K., Xiao, B. S., and Liu, I. L.B., 2014. Do actions speak louder than voices? The signaling role of social information cues in influencing consumer purchase decisions. *Decision Support Systems, 65*, pp.50–58.

Cho, H.J. and Pucik, V., 2005. Relationship between innovativeness, quality, growth, profitability, and market value. *Strategic Management Journal*, *26*(6), pp.555-575.

Claeys, A-S. and Cauberghe, V., 2014 What makes crisis response strategies work? The impact of crisis involvement and message framing. *Journal of Business Research*, *67*, pp.182–189.

Clementson, D. E., Pascual-Ferrá, P., & Beatty, M. J., 2016. How Language Can Influence Political Marketing Strategy and a Candidate's Image: Effect of Presidential Candidates' Language Intensity and Experience on College Students' Ratings of Source Credibility. *Journal of Political Marketing*, 15(4), pp.388–415.

Cole, S., Paulson, A. and Shastry, G.K., 2016. High school curriculum and financial outcomes: The impact of mandated personal finance and mathematics courses. *Journal of Human Resources*, *51*(3), pp.656-698.

Connelly, B. L., Certo, S. T., Ireland, R. D., and Reutzel, C. R. , 2011. Signaling Theory: A Review and Assessment. *Journal of Management, 37*(1), 39-67.

Courtney, C., Dutta, S., & Li, Y., 2017. Resolving information asymmetry: Signaling, endorsement, and crowdfunding success. *Entrepreneurship Theory and Practice*, *41*(2), pp.265-290.

Cumming, D. and Johan, S., 2017. The problems with and promise of entrepreneurial finance. *Strategic Entrepreneurship Journal*, *11*(3), pp.357-370.

Cumming, D.J., Johan, S.A. and Zhang, Y., 2019. The role of due diligence in crowdfunding platforms. *Journal of Banking & Finance*, *108*, pp.105661.

D'alcantara, G. and Gautier, A., 2013. The postal sector as a vector of financial inclusion. *Annals of Public and Cooperative Economics*, *84*(2), pp.119-137.

Danilov, A., & Sliwka, D., 2016. Can contracts signal social norms? Experimental evidence. *Management Science*, *63*(2), pp.459-476.

Davis, B. and Allison, T., 2013. Crowd Funding of Entrepreneurial Ventures: A Comparison of Signaling and Impression Management. *Academy of Management Proceedings*, 2013(1), p.16102.

Davis, B. C., Hmieleski, K. M., Webb, J. W., & Coombs, J. E., 2017. Funders' positive affective reactions to entrepreneurs' crowdfunding pitches: The influence of perceived product creativity and entrepreneurial passion. *Journal of Business Venturing*, *32*(1), pp.90-106.

Defazio, D., Franzoni, C., & Rossi-Lamastra, C., 2020. How Pro-social Framing Affects the Success of Crowdfunding Projects: The Role of Emphasis and Information Crowdedness. *Journal of Business Ethics*, pp.1–22.

Devlin, J.F., 2005. A detailed study of financial exclusion in the UK. *Journal of Consumer Policy*, *28*(1), pp.75-108.

Ding, J., Huang, J., Li, Y. and Meng, M., 2019. Is there an effective reputation mechanism in peer-to-peer lending? Evidence from China. *Finance Research Letters*, 30, pp.208-215.

Ding, R. & Hou, W., 2015. Retail investor attention and stock liquidity. *Journal of International Financial Markets, Institutions and Money,* 37**,** pp.12-26.

Diniz, E., Birochi, R. and Pozzebon, M., 2012. Triggers and barriers to financial inclusion: The use of ICT-based branchless banking in an Amazon county. *Electronic Commerce Research and Applications*, *11*(5), pp.484-494.

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

Dowell, D., Morrison, M. and Heffernan, T., 2015. The changing importance of affective trust and cognitive trust across the relationship lifecycle: A study of business-to-business relationships. *Industrial Marketing Management, 44*, pp.119–130.

Duan, T., Ding, R., Hou, W. & Zhang, J. Z., 2018. The burden of attention: CEO publicity and tax avoidance. *Journal of Business Research,* 87**,** pp.90-101.

Duarte, J., Siegel, S. and Young, L., 2012. Trust and Credit: The Role of Appearance in Peer-to-peer Lending. *The Review of Financial Studies, 25*(8), pp.2455–2484.

Dyer, J. H. and Chu, W., 2003. The Role of Trustworthiness in Reducing Transaction Costs and Improving Performance: Empirical Evidence from the United States, Japan, and Korea. *Organization Science, 14*(1), pp.57-68.

Economist. 2016. Credit in China: Just spend. *Economist*: Nov. 17.

Eisend, M. (2009). A meta-analysis of humor in advertising. *Journal of the Academy of Marketing Science,* 37**,** 191-203.

Elsbach, K. , 2004. Managing images of trustworthiness in organizations.

Engelberg, J. & Gao, P., 2011. In search of attention. *The Journal of Finance,* 66**,** pp.1461-1499.

Farrell, J., & Rabin, M. , 1996 . Cheap talk. *Journal of Economic Perspectives*, *10*(3), pp.103-118.

Fausey, C. M., & Boroditsky, L. , 2010. Subtle linguistic cues influence perceived blame and financial liability. *Psychonomic Bulletin & Review*, 17(5), pp.644–650.

Fernandes, D., Lynch Jr, J.G. and Netemeyer, R.G., 2014. Financial literacy, financial education, and downstream financial behaviors. *Management Science*, *60*(8), pp.1861-1883.

Flanagin, A.J., 2007. Commercial markets as communication markets: Uncertainty reduction through mediated information exchange in online auctions. *New Media & Society*, *9*(3), pp.401-423.

Freedman, S., & Jin, G. Z., 2017. The information value of online social networks: lessons from peer-to-peer lending. *International Journal of Industrial Organization*, *51,* pp.185-222.

Freedman, S.M. and Jin, G.Z., 2011. *Learning by doing with asymmetric information: Evidence from Prosper.com* (No. w16855). National Bureau of Economic Research.

French, D. and McKillop, D., 2016. Financial literacy and over-indebtedness in low-income households. *International Review of Financial Analysis*, *48,* pp.1-11.

Fukuyama, F., 1995. *Trust*. New York: Free Press.

Fungacova, Z. and Weill, L., 2014. A view on financial inclusion in Asian countries. *BOFIT Policy Brief*, *2014*(8), p.3.

Fuster Wang, J., 2018. Inclusion or expulsion: Digital technologies and the new power relations in China's "Internet finance". *Communication and the Public*, *3*(1), pp.34-45.

Gao, H., Yu, T., & Cannella Jr, A. A., 2016. The use of public language in strategy: A multidisciplinary review and research agenda. *Journal of Management*, *42*(1), pp.21-54.

García, M. N., Muñoz-Gallego, P. A., Viglia, G., & González-Benito, Ó., 2020. Be Social! The Impact of Self-Presentation on Peer-to-Peer Accommodation Revenue: *Journal of Travel Research*, 59(7), pp.1268–1281.

Giora, R., 2006. Anything negatives can do affirmatives can do just as well, except for some metaphors. *Journal of Pragmatics*, 38(7), pp.981–1014.

Grant, C., 2007. Estimating credit constraints among US households. *Oxford Economic Papers*, *59*(4), pp.583-605.

Guo, M., Liao, X., Liu, J. and Zhang, Q., 2019. Consumer preference analysis: A data-driven multiple criteria approach integrating online information. *Omega*, p.102074.

Guo, Y., Zhoub, W., Luo, C., Liu, C. and Xiong, H., 2016. Instance-based credit risk assessment for investment decisions in P2P lending. *European Journal of Operational Research 249*, pp.417–426.

Hamilton, B.H. and Nickerson, J.A., 2003. Correcting for endogeneity in strategic management research. *Strategic Organization*, *1*(1), pp.51-78.

Han, J.T., Chen, Q., Liu, J.G., Luo, X.L. and Fan, W., 2018. The persuasion of borrowers' voluntary information in peer to peer lending: An empirical study based on elaboration likelihood model. *Computers in Human Behavior*, *78*, pp.200-214.

Han, S., Lerner, J. S., & Keltner, D., 2007. Feelings and Consumer Decision Making: The Appraisal-Tendency Framework. *Journal of Consumer Psychology*, 17(3), pp.158–168.

Harmon, D. J., Green Jr, S. E., & Goodnight, G. T., 2015. A model of rhetorical legitimation: The structure of communication and cognition underlying institutional maintenance and change. *Academy of Management Review*, *40*(1), pp.76-95.

Harms, M., 2007. What Drives Motivation to Participate Financially in a Crowdfunding Community?. *SSRN Electronic Journal*,.

Hasan, A., Moin, S., Karim, A. and Shamshirband, S., 2018. Machine learning-based sentiment analysis for twitter accounts. *Mathematical and Computational Applications*, *23*(1), p.11.

Hasan, M.M., Yajuan, L. & Mahmud, A. 2020, Regional Development of China's Inclusive Finance Through Financial Technology, *SAGE Open*, 10(1).

Heckman, J.J., 1979. Sample selection as a specification error. *Econometrica*, *47*, pp.153-161.

Herzenstein, M., Andrews, R. L., Dholakia, U. M., & Lyandres, E. , 2008. The democratization of personal consumer loans? Determinants of success in online peer-to-peer lending communities. *Boston University School of Management Research Paper*, *14*(6).

Herzenstein, M., Sonenshein, S. & Dholakia, U. M., 2011. Tell me a good story and I may lend you money: The role of narratives in peer-to-peer lending decisions. *Journal of Marketing Research*, 48**,** pp.138-149.

Hoefer, R. L., & Green Jr, S. E., 2016. A rhetorical model of institutional decision making: The role of rhetoric in the formation and change of legitimacy judgments. *Academy of Management Review*, *41*(1), pp.130-150.

Homer, P. M. and Yoon, S. G., 1992. Message Framing and the Interrelationships Among Ad-Based Feelings, Affect and Cognition. *Journal of Advertising*, *21*(1), pp.19-33.

Honohan, P., 2008. Cross-country variation in household access to financial services. *Journal of Banking & Finance*, *32*(11), pp.2493-2500.

HOU, Y., MA, X., MEI, G., WANG, N. and XU, W., 2019. A Trial of Student Self-Sponsored Peer-to-Peer Lending Based on Credit Evaluation Using Big Data Analysis. *Computational Intelligence and Neuroscience*.

Howell, N. and Wilson, T., 2005. Access to consumer credit: The problem of financial exclusion in Australia and the current regulatory framework. *Macquarie LJ*, *5*, pp.127.

Hsu, P.Y., Lei, H.T., Huang, S.H., Liao, T.H., Lo, Y.C. and Lo, C.C., 2019. Effects of sentiment on recommendations in social network. *Electronic Markets*, *29*(2), pp.253-262.

Huang, R. H., 2018. Online P2P Lending and Regulatory Responses in China: Opportunities and Challenges. *European Business Organization Law Review*, *19*(1), pp.63-92.

Huang, W., Qian, Y., & Xu, N., 2020. The signaling effects of education in the online lending market: Evidence from china. *Economic Modelling*

Hui, X.M., 2010. Subjective Meanings of Negation. *Foreign Language Research*, 6, pp. 6–8.

Humphreys, A., & Wang, R. J. H., 2018. Automated Text Analysis for Consumer Research. *Journal of Consumer Research*, 44(6), pp.1274–1306.

Huston, S.J., 2015. Using a Financial Health Model to Provide Context for Financial Literacy Education Research: A Commentary. *Journal of Financial Counseling and Planning*, *26*(1), pp.102-104.

Iacobucci, D., Saldanha, N. and Deng, X., 2007. A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of consumer psychology*, *17*(2), pp.139-153.

Ibtissem, B. and Bouri, A., 2013. Credit risk management in microfinance: The conceptual framework. *ACRN Journal of Finance and Risk Perspectives*, *2*(1), pp.9-24.

Jensen, M. L., Averbeck, J. M., Zhang, Z., & Wright, K. B., 2013. Credibility of Anonymous Online Product Reviews: A Language Expectancy Perspective. *Journal of Management Information Systems*, 30(1), pp.293–324.

Jensen, M. L., Averbeck, J. M., Zhang, Z., & Wright, K. B., 2013. Credibility of Anonymous Online Product Reviews: A Language Expectancy Perspective. *Journal of Management Information Systems*, 30(1), pp.293–324.

Jiang, X., Li, Y., & Zhou, X., 2013. Is it over-respectful or disrespectful? Differential patterns of brain activity in perceiving pragmatic violation of social status information during utterance comprehension. *Neuropsychologia*, 51(11), pp.2210–2223.

Johnson, D. and Grayson, K., 2005. Cognitive and affective trust in service relationships. *Journal of Business research*, *58*(4), pp.500-507.

Johnston Jr, D. and Morduch, J., 2008. The unbanked: evidence from Indonesia. *The World Bank Economic Review*, *22*(3), pp.517-537.

Kaplan, S. , & Vakili, K., 2014. The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal*, *36*(10), pp.1435-1457.

Ke, R. and Zhang, W, 2003. Trust in China: A cross-regional analysis. William Davidson Institute Working Paper Number 586, William Davidson Institute, The University of Michigan Business School.

Kempson, E., Whyley, C., 1999a. *Kept in or Opted Out? Understanding and Combating Financial Exclusion*. Policy Press, Bristol.

Kempson, E., Whyley, C., 1999b. *Understanding and combating financial exclusion*. Insurance Trends (Assoc. Br. Insurers), 21, 18–22

Koku, P.S., 2015. Financial exclusion of the poor: A literature review. *International Journal of Bank Marketing*

Komarova Loureiro, Yuliya and Laura Gonzalez (2015),"Competition against Common Sense: Insights on Peer-to-Peer Lending as a Tool to Allay Financial Exclusion," *International Journal of Bank Marketing (Special Issue on Financial Exclusion)*

Kromidha, E., & Robson, P., 2016. Social identity and signaling success factors in online crowdfunding. *Entrepreneurship & Regional Development*, *28*(9-10), pp.605-629.

Krugman, D. M., Fox, R. J., Fletcher, J. E., Fischer, P. M. & Rojas, T. H., 1994. Do adolescents attend to warnings in cigarette advertising? An eye-tracking approach. *Journal of Advertising Research,* 34**,** pp.39-39.

Landström, H. (2017). *Advanced Introduction to Entrepreneurial Finance*. Edward Elgar Publishing.

Larrimore, L., Jiang, L., Larrimore, J., Markowitz, D., & Gorski, S., 2011. Peer to Peer Lending: The Relationship Between Language Features, Trustworthiness, and Persuasion Success. *Journal of Applied Communication Research*, 39(1), pp.19–37.

Lee, C. H., Bian, Y., Karaouzene, R., & Suleiman, N. (2019). Examining the role of narratives in civic crowdfunding: linguistic style and message substance. *Industrial Management and Data Systems*.

Lee, H-H. M. and van Dolen, W., 2015. Creative participation: Collective sentiment in online co-creation communities. *Information & Management, 52*, pp.951–964.

Lee, R.P., Chen, Q. and Hartmann, N. N., 2016. Enhancing Stock Market Return with New Product Preannouncements: The Role of Information Quality and Innovativeness, *Journal of Product Innovation Management*, *33*(4), pp.455-471.

Li, J. & Zhan, L., 2011. Online Persuasion: How the Written Word Drives WOM: Evidence from Consumer-Generated Product Reviews. *Journal of Advertising Research*, *51*(1), pp. 239-57.

Li, J., & Hu, J., 2019. Does university reputation matter? Evidence from peer-to-peer lending. *Finance Research Letters,* 31, pp.66-77

Li, T., Wei X.P., Zheng Z.X., Yi X.J., Zhao X.R., He X.Y.(2019). The effects of different power relations on negation bias of negative descriptions. *Acta Psychologica Sinica*, 51(6), 714-723.

Liao, L., Wang, Z., Xiang, J., Yan, H. and Yang, J., 2019. Investing with Fast Thinking. *Kelley School of Business Research Paper*, (18-54).

Lin, J.-., 2003, Aspectual selection and negation in Mandarin Chinese, *Linguistics*, 41(3), pp. 425-459.

Lin, M. and Viswanathan, S., 2015. Home bias in online investments: An empirical study of an online crowdfunding market. *Management Science*, *62*(5), pp.1393-1414.

Lin, M., Prabhala, N. R., & Viswanathan, S., 2013. Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science*, *59*(1), pp.17-35.

Liu, D., Brass, D. J., Lu, Y. E., & Chen, D., 2015. Friendships in Online Peer-to-Peer Lending: Pipes, Prisms, and Relational Herding. *MIS Quarterly, 39*(3), pp.729-742.

Liu, Y., Zhou, Q., Zhao, X. and Wang, Y., 2018. Can listing information indicate borrower credit risk in online peer-to-peer lending?. *Emerging Markets Finance and Trade*, *54*(13), pp.2982-2994.

Loewenstein, G., Sunstein, C.R. and Golman, R., 2014. Disclosure: Psychology changes everything. *Annual Review of Economics*, *6*(1), pp.391-419.

Ludwig, S., Ruyter, K. de, Friedman, M., Brüggen, E. C., Wetzels, M., & Pfann, G., 2013. More Than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates. *Journal of Marketing*, 77(1), pp.87–103.

Lundberg, U., Bratfisch, O., & Ekman, G., 1972. Emotional Involvement and Subjective Distance: A Summary of Investigations. *Journal of Social Psychology*, 87(2), pp.169–177.

Luthans, F., Luthans, K. W., & Luthans, B. C., 2004. Positive psychological capital: Beyond human and social capital. *Business Horizons*, *47*(1), pp.45-50.

Maglio, S. J., 2020. Psychological distance in consumer psychology: Consequences and antecedents. *Consumer Psychology Review*, 3(1), pp.108–125.

Maheswaran, D. and Meyers-Levy, J., 1990. The Influence of Message Framing and Issue Involvement. *Journal of Marketing Research*, *27*(3), pp.361-367.

Majumdar, A., & Bose, I., 2018. My words for your pizza: An analysis of persuasive narratives in online crowdfunding. Information & Management, 55(6), pp.781–794.

Martens, M. L., Jennings, J. E. & Jennings, P. D., 2007. Do the stories they tell get them the money they need? The role of entrepreneurial narratives in resource acquisition. *Academy of Management Journal*, 50**,** pp.1107-1132.

Martí, J. and Balboa, M., 2007. Characterization of the reputation of private equity managers: Evidence in Spain. *Journal of Business Venturing*, *22*(4), pp.453-480.

Matson, E. 1999, Capital rationing, *Scandinavian Journal of Management*, 15(2), pp. 157-171.

McAllister, D.J., 1995. Affect and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Review*, 20, pp.709–34.

McKenny, A. F., Allison, T. H., Ketchen Jr, D. J., Short, J. C., & Ireland, R. D., 2017. How should crowdfunding research evolve? A survey of the entrepreneurship theory and practice editorial board. *Entrepreneurship Theory and Practice*, *41*(2), pp.291-304.

McKillop, D. and Wilson, J., 2007. Financial exclusion. *Public Money and Management*, *27*(1), pp.9-12.

Mehl, M. R., Robbins, M. L., & Holleran, S. E., 2013. How Taking a Word for a Word Can Be Problematic: Context-Dependent Linguistic Markers of Extraversion and Neuroticism. Journal of Methods and Measurement in the Social Sciences, 3(2), pp.30–50.

Mi, J.J. and Zhu, H., 2017. Can funding platforms' self-initiated financial innovation improve credit availability? Evidence from China's P2P market. *Applied Economics Letters*, *24*(6), pp.396-398.

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

Miller, G., 1995. *The science of words*. New York: Scientific American Library.

Miller, M., Reichelstein, J., Salas, C. and Zia, B., 2014. *Can you help someone become financially capable? A meta-analysis of the literature*. The World Bank.

Mills, K. and McCarthy, B., 2016. The state of small business lending: Innovation and technology and the implications for regulation. *Harvard Business School Entrepreneurial Management Working Paper*, (17-042), pp.17-042.

Milne, A. and Parboteeah, P. (2016). The business models and economics of peer-to-peer lending.

Mittal, S., & Lloyd, J., 2016. *The rise of FinTech in China: redefining financial services. A Collaborative Report.* DBS and Ernst & Young (EY). https://www.ey.com/Publication/vwLUAssets/ey-the-rise-of-fintech-in-china/$FILE/ey-the-rise-of-fintech-in-china.pdf (Accessed 4 January 2019).

Newman, A., Ucbasaran, D., Zhu, F. E. I., & Hirst, G., 2014. Psychological capital: A review and synthesis. *Journal of Organizational Behavior*, *35*(S1), pp. S120-S138.

Nook, E.C., Schleider, J.L. & Somerville, L.H. 2017, A linguistic signature of psychological distancing in emotion regulation, *Journal of Experimental Psychology: General*, 146(3), pp. 337-346.

Norton, E. C., Wang, H., & Ai, C., 2004. Computing interaction effects and standard errors in logit and probit models. *The Stata Journal*, 4(2), pp.154-167.

Office of Fair Trading, 1999. *Vulnerable customers and financial services*. Rep. No. OFT 255. Office of Fair Trading. London

Öhman, A., Flykt, A., & Esteves, F., 2001. Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), pp.466.

ONAY, C. and ÖZTÜRK, E., 2018. A review of credit scoring research in the age of Big Data. *Journal of Financial Regulation and Compliance*, 26(3), pp. 382-405.

Ozili, P.K., 2020. Theories of Financial Inclusion. *SSRN Electronic Journal* .

Packard, Grant, Sarah G. Moore, and Brent McFerran, 2014, How Can "I" Help "You"? The Impact of Personal Pronoun Use in Customer-Firm Agent Interactions, *MSI Report*, pp.14-110.

Palmie, M., Wincent, J., Parida, V. & Caglar, U. 2020, The evolution of the financial technology ecosystem: An introduction and agenda for future research on disruptive innovations in ecosystems, *Technological Forecasting and Social Change*, 151.

Panos, G.A. & Wilson, J.O.S. 2020, Financial literacy and responsible finance in the FinTech era: capabilities and challenges, *European Journal of Finance*, 26(4-5), pp. 297-301.

Parhankangas, A., & Renko, M. , 2017. Linguistic style and crowdfunding success among social and commercial entrepreneurs. *Journal of Business Venturing*, 32(2), pp.215–236.

Peng, K., Fan, Y., & Hsu, T., 2004. Proposing the content perception theory for the online content industry – a structural equation modeling. *Industrial Management and Data Systems*, 104(6), pp.469–489.

Pennebaker, J. W., & King, L. A., 1999. Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77, pp.1296–1312.

Pennebaker, J. W., 2011. The Secret Life of Pronouns: What Our Words Say About Us.

Pennebaker, J., & Chung, C. K., 2013. *Counting little words in big data: The psychology of individuals, communities, culture, and history*, 25–42.

Pennebaker, James W., 2011. The Secret Life of Pronouns, *New Scientist*, 211 (2828), pp.42-45.

Pham, T., & Talavera, O., 2018. Discrimination, social capital, and financial constraints: The case of Viet Nam. *World Development*, *102*, pp.228-242.

Pieters, R. & Wedel, M., 2004. Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing*, 68**,** pp.36-50.
Plummer, L. A., Allison, T. H., & Connelly, B. L., 2016. Better together? Signaling interactions in new venture pursuit of initial external capital. *Academy of Management Journal*, *59*(5), pp.1585-1604.
Preacher, K.J., Rucker, D.D. and Hayes, A.F., 2007. Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate behavioral research*, 42(1), pp.185-227.
Putnam, R., 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.
Revenstorf, D., Hahlweg, K., Schindler, L. and Vogel, B., 1984. Interaction analysis of marital conflict. *Marital interaction: Analysis and modification*, pp.159-181.

Robinson, R.L., Navea, R., Ickes, W., 2013. *Predicting final course performance from students' written self-introductions: a LIWC analysis*. J. Lang. Soc. Psychol. XX, 1–11

Saiedi, E., Mohammadi, A., Broström, A., & Shafi, K. (2020). Distrust in Banks and Fintech Participation: The Case of Peer-to-Peer Lending: *Entrepreneurship Theory and Practice*.
Salehan, M., & Kim, D.J., 2016. Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems*, 81, pp.30–40.
Salomon, V., 2018. Strategies of startup evaluation on crowd-investing platforms: The case of Switzerland. *Journal of Innovation Economics Management*, (2), pp.63-88.
Schweitzer, M.E. and Barkley, B., 2017. Is' Fintech'Good for Small Business Borrowers? Impacts on Firm Growth and Customer Satisfaction.

Shaver, J.M., 2005. Testing for mediating variables in management research: Concerns, implications, and alternative strategies. *Journal of management*, *31*(3), pp.330-353.
Simmons, R.A., Gordon, P.C. and Chambless, D.L., 2005. Pronouns in marital interaction: What do “you” and “I” say about marital health?. *Psychological science*, *16*(12), pp.932-936.

Solo, T.M., 2008. Financial exclusion in Latin America—or the social costs of not banking the urban poor. *Environment and Urbanization*, *20*(1), pp.47-66.

Sparreboom, P. and Duflos, E., 2012. Financial inclusion in the People’s Republic of China: An analysis of existing research and public data. *China Papers on Inclusiveness*, *7*, pp.11-64.

Spence, M., 1973. Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355-374.
Steigenberger, N., & Wilhelm, H., 2018. Extending Signaling Theory to Rhetorical Signals: Evidence from Crowdfunding. *Organization Science*, *29*(3), pp.529-546.
SWAIDAN, Z., VITELL, S.J. and RAWWAS, M.Y.A., 2003. Consumer Ethics: Determinants of Ethical Beliefs of African Americans. *Journal of Business Ethics*, **46**(2), pp. 175-186.

Tan, H. H., & Chee, D., 2005. Understanding interpersonal trust in a Confucian-influenced society: An exploratory study. *International Journal of Cross Cultural Management*, *5*(2), pp.197-212.
Tan, L, 2014. Problems in Development of Micro-credit Loans to Farmers in China and Countermeasures, *Asian Agricultural Research, USA-China Science and Culture Media Corporation*, vol. 6(5), pp. 1-4, May.
Tao, Q., Dong, Y. and Lin, Z., 2017. Who can get money? Evidence from the Chinese peer-to-peer lending platform. *Information Systems Frontiers*, *19*(3), pp.425-441.

Tausczik, Y. R., & Pennebaker, J. W. , 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), pp.24–54.

Tidwell, L.C. and Walther, J.B., 2002. Computer‐mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human Communication Research*, *28*(3), pp.317-348.

Tobias Berg & Valentin Burg & Ana Gombović & Manju Puri & Andrew Karolyi, 2020. On the Rise of FinTechs: Credit Scoring Using Digital Footprints, *The Review of Financial Studies*, 33(7), pp. 2845-2897.

Toma, C.L. & Hancock, J.T. 2012, What Lies Beneath: The Linguistic Traces of Deception in Online Dating Profiles, *Journal of Communication*, 62(1), pp. 78-97.

Toma, C.L., D'Angelo, J.D., 2015. Tell-tale words: linguistic cues used to infer the expertise of online medical advice. *J. Lang. Soc. Psychol*. 34, pp. 25–45

Wang, H., Choi, J., & Li, J., 2008. Too little or too much? Untangling the relationship between corporate philanthropy and firm financial performance. *Organization Science*, *19*(1), pp.143-159.

Wang, J., Shen, Y. and Huang, Y., 2016. Evaluating the regulatory scheme for internet finance in China: the case of peer-to-peer lending. *China Economic Journal*, *9*(3), pp.272-287.

Welter, F., & Smallbone, D., 2006. Exploring the role of trust in entrepreneurial activity. *Entrepreneurship Theory and Practice*, *30*(4), pp.465-475.

Williams, L. E., Stein, R., & Galguera, L., 2014. The Distinct Affective Consequences of Psychological Distance and Construal Level. *Journal of Consumer Research*, 40(6), pp.1123–1138.

Williams, L. E., Stein, R., & Galguera, L., 2014. The distinct affective consequences of psychological distance and construal level. *Journal of Consumer Research*, *40*(6), pp.1123-1138.

Wilson, T., Wiebe, J. and Hoffmann, P., (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*.

Wilson, T.A., 2012. Supporting social enterprises to support vulnerable consumers: the example of community development finance institutions and financial exclusion. *Journal of consumer policy*, *35*(2), pp.197-213.

World Bank, 2014. *Financial Inclusion*, Global Financial Development Report.

Xiao, R. & McEnery, T. 2008, Negation in Chinese: A Corpus-based study, *Journal of Chinese Linguistics*, 36(2), pp. 274-320.

Xu, J. J., & Chau, M., 2018. Cheap talk? The impact of lender-borrower communication on peer-to-peer lending outcomes. *Journal of Management Information Systems*, *35*(1), pp. 53-85.

Xu, J., 2017. China′s Internet Finance： A Critical Review. *China & World Economy*, 25(4), pp.78–92.

Xu, J., Hilliard, J., & Barth, J. R., 2018. On education level and terms in obtaining P2P funding: New evidence from china. *International Review of Finance*

Xu, Q., Fernando, G.D. and Tam, K., 2019. Trust and firm performance: A bi-directional study. *Advances in Accounting*, *47*, p.100433.

Xu, T.J.(2015). Negation Bias of Linguistic and Its Effect on the impression Valuation (Master Dissertation). Huazhong Normal University, Wuhan

YAN, J., WANG, K., LIU, Y., XU, K., KANG, L., CHEN, X. and ZHU, H., 2018. Mining social lending motivations for loan project recommendations. *Expert Systems with Applications*, **111**, pp. 100-106.

Yang, C., 2007. Chinese Internet language: A sociolinguistic analysis of adaptations of the Chinese writing system. *language@ internet, 4*(2).

Yao, J., Chen, J., Wei, J., Chen, Y. & Yang, S., 2018.The relationship between soft information in loan titles and online peer-to-peer lending: evidence from RenRenDai platform. *Electronic Commerce Research***,** pp.1-19.

Zhang, L., Li, X.P., Zhang, F.B. and Hu, B., 2017. Research on Keyword Extraction and Sentiment Orientation Analysis of Educational Texts. *Journal of Computers*, *28*(6), pp.301-313.

Zhang, S.T., Wang, F.F., Duo, F. and Zhang, J.L., 2018. Research on the Majority Decision Algorithm based on WeChat sentiment classification. *Journal of Intelligent & Fuzzy Systems*, *35*(3), pp.2975-2984.

Zhang, Y. and Buda, R., 1999. Moderating Effects of Need for Cognition on Responses to Positively versus Negatively Framed Advertising Messages**.** *Journal of Advertising, 28*(2), pp.1-15.

Zhao, X., Lynch Jr, J.G. and Chen, Q., 2010. Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37(2), pp.197-206.

Zhao, Y. P., Hu, P and Ma, Y. F., 2016. The issues of P2P platform in-house credit rating and solutions---The cases of Renrendai and Paipaidai. *Hainan Finance*, (6), pp.65-69. In Chinese

Zhu, L.Y. and Meeks, C.B., 1994. Effects of low income families' ability and willingness to use consumer credit on subsequent outstanding credit balances. *Journal of Consumer Affairs*, *28*(2), pp.403-422.

Zuo, W., Zhu, W., Chen, S. and He, X., 2019. Service quality management of online car-hailing based on PCN in the sharing economy. *Electronic Commerce Research and Applications*, *34*, p.100827.