

Wang, Yao; Drabek, Zdenek; Wang, Zhengwei

Working Paper

The predicting power of soft information on defaults in the Chinese P2P lending market

IES Working Paper, No. 20/2018

Provided in Cooperation with:

Charles University, Institute of Economic Studies (IES)

Suggested Citation: Wang, Yao; Drabek, Zdenek; Wang, Zhengwei (2018) : The predicting power of soft information on defaults in the Chinese P2P lending market, IES Working Paper, No. 20/2018, Charles University in Prague, Institute of Economic Studies (IES), Prague

This Version is available at:

<https://hdl.handle.net/10419/203198>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Institute of Economic Studies, Faculty of Social Sciences
Charles University in Prague

The Predicting Power of Soft Information on Defaults in the Chinese P2P Lending Market

Yao Wang
Zdenek Drabek
Zhengwei Wang

IES Working Paper: 20/2018



Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
<http://ies.fsv.cuni.cz>

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00 Praha 1

E-mail : ies@fsv.cuni.cz
<http://ies.fsv.cuni.cz>

Disclaimer: The IES Working Papers is an online paper series for works by the faculty and students of the Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Czech Republic. The papers are peer reviewed. The views expressed in documents served by this site do not reflect the views of the IES or any other Charles University Department. They are the sole property of the respective authors. Additional info at: ies@fsv.cuni.cz

Copyright Notice: Although all documents published by the IES are provided without charge, they are licensed for personal, academic or educational use. All rights are reserved by the authors.

Citations: All references to documents served by this site must be appropriately cited.

Bibliographic information:

Wang Y., Drabek Z. and Wang Z. (2018): "The Predicting Power of Soft Information on Defaults in the Chinese P2P Lending Market" IES Working Papers 20/2018. IES FSV. Charles University.

This paper can be downloaded at: <http://ies.fsv.cuni.cz>

The Predicting Power of Soft Information on Defaults in the Chinese P2P Lending Market

Yao Wang^a

Zdenek Drabek^a

Zhengwei Wang^b

^aInstitute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Smetanovo nábřeží 6, 111 01 Prague 1, Czech Republic

^bTsinghua University, PBC School of Finance

Email (corresponding author): zdenek.drabek1@gmail.com

August 2018

Abstract:

Online peer to peer lending (P2P)– allows people who want to borrow money to submit their applications on the platform and individual investors can make bids on the loan listings. The quality of information in credit appraisal becomes paramount in this market. The existing research to assess the role of what is known as soft information in P2P markets has so far been very limited and, inconclusive due to differences in approaches and methodological limitations. The aim of the paper is to discuss the role of soft information channels in predicting defaults in the P2P lending market and to assess the importance of soft information in the Fintech companies' credit analysis. Using a unique data of the Chinese P2P lending platform RRDai.com and new approach based on sets of hard and soft information, we compare the predicting performance of soft information, hard information and the combined role of both hard and soft information. We show that soft information can provide a valuable input in credit appraisal. The predicting power of soft information in our test was high, and together with hard information it can even help improve the loan performance. In exceptional situations characterized by the absence of hard financial data, soft information could be used, with caution, as an alternative.

JEL: D82; E51; G02;G14;G21;G23

Keywords: Soft Information, P2P Lending, Fintech, Microfinance, Credit Analysis, Empirical Study

Acknowledgements: In the preparation of the paper Laura has benefited from numerous discussions with the staff and graduate students of IES at Charles University. We are also grateful dr. Ira Lieberman, former lead economist at the World Bank, for his insights and suggestions. Comments of an anonymous referee for WP IES of Charles University are gratefully acknowledged. None of these persons should be seen as responsible for any remaining shortcomings.

1 Introduction

The microfinance industry has undergone fairly dramatic changes over the last few decades. Perhaps one of the most interesting examples is the development of peer-to-peer (P2P) lending and the utilization of new technologies which have spectacularly transformed the industry in some countries these days. The first P2P platform Zopa started in the UK in 2005, followed by Prosper and Lending Club in 2006. Then in 2007, the P2P platforms emerged in other European countries (e.g., Smava in Germany, TrustBuddy in Sweden, Prestiamoci in Italy), China (e.g., PPDai), Japan (e.g., Maneo). Since 2009, the P2P platforms have been booming on the global scale.

Microfinance, an alternative source of borrowing by SME and individuals, has also seen a vast growth¹, but it is costly to intermediate. The size of loans is small, and it requires proximity of lenders to borrowers and hence the presence of branches and loan offices working in villages or local communities. Moreover, the microfinance sector has been facing two major challenges over the last decade or so: One is digital (mobile) banking and the other is Fintech of which P2P lending is a sub-set. The establishment of P2P platforms has allowed the lowering of transaction costs for both lenders and borrowers through disintermediation. New information technologies and better information have provided more room to reduce uncertainties and for greater efficiencies of lending, opportunities to expand markets, and to improve access to credit.

The quality and quantity of information available to lenders and borrowers have been subject to particularly significant changes in those markets. Traditionally, the industry has relied on financial statements as the predictor of credit standing (Horrigan,1966). However, after decades, various researchers have identified a number of shortcomings of the system which is based purely on financial statements. For example, Jiménez and Saurina (2004) found that collateral could not secure repayment of loans and that loans with collateral sometimes actually have higher default rates. Berger *et al.* (2005) show that a credit scoring system contributes to increasing the credit availability for small businesses, but its function for credit risk analysis is not as effective as expected. A similar result has been obtained by DeYoung *et al.* (2008), who found a positive relationship between the use of a credit scoring model and the loan default rates.

¹ The SME sector was serving 10 million clients in 1995, the early days of SME lending. And the corresponding figure today is around 300 million. Microfinance institutions have been commercialized over time and today around 100 specialized funds have invested and loaned about US\$ 12.5 billion.

To fix the flaws of these traditional - hard - information based credit rationing systems, the use of soft information became a possible alternative and complementary solution. A great number of studies have now been conducted on bank relationship lending - a lending method that is prevalent among small banks in SME lending. It is based on the idea that a relationship with the borrower provides “soft” information such as the reliability of the business owner. This soft information is collected through long term contacts with the small enterprises and their local communities. It is valuable but time-consuming to collect, and it is tough to convey.

With technological development, the availability of social and psychological soft information is increasing while the costs of collecting information are decreasing. The body of research drawing on soft information data has been growing. Interesting work has been done on the role of race in lending (e.g., Ravina, 2012), appearance (e.g., Gonzalez & Loureiro, 2014), age (e.g., Pope Sydnor, 2011), gender (e.g., Barasinska, 2010) and even linguistic factors (e.g., Dorfleitner *et al.* 2016). However, research focusing on soft information as a whole is still limited and inconclusive. Methodologies and approaches of studying the role of “soft” info have been different driven by different questions of researchers. Most of the research has been focused on determinants of loans without enough attention being paid to the efficiency of lending. Existing research typically deals with a particular market or a particular factor of lending. By far the most attention has been paid to SME markets and very little empirical work has been done targeting the role of soft information on P2P platforms. Surprisingly, virtually no research has been done on perhaps one of the most buoyant P2P market in the world – China. Only few empirical studies exist and they have covered different markets. The one conclusion that can be drawn from those studies is that markets are not homogeneous and are most likely made up by differences in institutional structures, cultures and psychological factors.

The aim of the paper is to discuss the role of different of information channels in predicting defaults in the P2P lending market and to assess the importance of soft information in the Fintech companies’ credit analysis. We hope to answer the following question - can risk assessment be facilitated by the incorporation of soft information in the appraisal of credit risk? To do so, we have built a model of determinants of loan defaults which looks at the importance of soft and hard information in different scenarios. The paper compares the predicting performance of soft information, hard information and the combined role of both hard and soft information. The results show that soft information can provide a valuable input in credit appraisal. The predicting power of soft information in our test was high, and together with hard information it could help improve the loan performance.

The paper is divided into six sections. The following Section 2 provides a review of the relevant empirical literature. Its purpose is to identify the important advances in the debate about the quality of information and the key gaps and limitations of the literature which will drive our approach and methodology. Section 3 describes the methodology: the data used in the study; and the econometric method – logit model – with which we are evaluating the role of soft information in credit appraisal. The results of our empirical test are presented in Section 4. We have also carried sensitivity tests of our estimations, and the results are reported in Section 5. Our conclusions are summarized in the final Section 6.

2 Literature Review

2.1 Foundations of Soft Information in Credit Appraisal

The quality and availability of the right information are as critical in the P2P market as they are in the whole financial industry.² Asymmetric information, adverse selection, and moral hazard are all issues as important in P2P platforms as they are for loan officers in big banks. It has been common in the financial industry in general, and in the P2P markets especially, for lenders to rely on financial statements. However, following the work of Spence (1973), attention has increasingly been paid to other information that may “signal” the ability and willingness of borrowers to repay loans. Spence himself devoted a great deal of attention to signals obtained from the educational attainments of clients, which provided a signaling function to creditors/investors in their assessments to differentiate among clients in terms of their credit status.³

² The SME lending institutions have developed in multiple forms. The original form were NGOs which were subsequently joined by cooperatives, credit unions, non-bank financial institutions and microfinance banks.

³ Following Ashta & Assadi (2009) P2P lending can be divided into commercial and non-commercial. The main difference between the two concepts is the treatment of credit risk. While borrowers are charged interest rate according to their risks on the commercial P2P lending platform, this is not the case on non-commercial P2P platforms. The non-commercial platforms are basically a mode of social fund raising, typically among friends or people of similar persuasion, interests or objectives. This paper deals entirely with the former category of P2P platforms – commercial lending.

Over time, researchers have distinguished between two categories of information. Jensen and Meckling (1992) divided the knowledge of agents into “specific and general knowledge”. They defined the “specific” knowledge as the knowledge that is costly to collect while “general” knowledge which is relatively cheaper to transfer. This is analogous with hard information. Based on this concept and using the framework of organizational structure for the analysis, they investigated the relationship between costly “specific” information and decentralization and concluded that decisions should always be made by those agents who possess “specific” information on efficiency grounds.

The distinction between soft and hard information can be traced back to the Diamond (1984)’s theory of financial intermediaries. In distinguishing between banks and the public bond market, Diamond emphasized the different abilities of agents in the two markets to collect and process information. With the help of rating agencies, public market players can obtain the necessary “hard” information such as financial statements, credit histories or default records. However, there is one element missing for bond investors – the direct contact with borrowers. This source of data is precisely the element of credit appraisal that banks can rely on when they have direct contacts with their clients. By focusing on so-called relationship lending, banks can obtain “soft” information about their clients, which has been particularly important in micro-finance. As Berger and Udell (2002) showed, small business loans rely more on relationship lending due to a lack of “hard” information. This is particularly the case for small banks. Large banks are hierarchical and, thus, rely more on the transferable “hard” information while small banks are more efficient in generating and processing “soft” information based relationship lending (Stein, 2002).⁴

Later on, Petersen (2004) framed the distinction between soft and hard information in terms of characteristic of each category, their advantages and disadvantages and in terms of the process of hardening soft information. The main characteristics of “hard” information are recorded as numbers, comparable, easy to collect, store and convey. He also noted that due to their numerical feature, hard information constitutes the basis for big data programs in credit appraisals. In contrast, soft information was usually presented in texts which include opinions, ideas, business plans or even rumors.

. In current practices, hard credit information refers to information such as credit rating or all the registered information from the central bank or any credit reference system, like the FICO

⁴ By way of example, credit scores are very important in the US Community Banking Sector even though soft information is increasingly utilized.

score from the US and the credit report from credit reference center of PBOC in China. It contains the borrower's key financial information like income, mortgage loans, and debt to income ratio, property ownership, credit record and such. These are the standard references that are used by traditional banks to analyze the creditworthiness of a borrower. Soft information refers to information that is more difficult to quantify. It contains information based on judgments about the personality and reliability of the borrower, the state of the business seeking a loan, the potential for future development, etc. Soft information may even include a borrower's social networks, video interviews, profile picture, description of the borrowing stories, etc. The collection of "soft" information is a cumulative process; it comes from long-term contact with firm owners and other individuals who are seeking loans.

Empirical evidence suggests that the traditional lending method based on credit scoring system and hard information links to riskier loans with higher price, thus it may sometimes not be a reliable predictor of default (Berger et al., 2005 and DeYoung et al., 2008). Furthermore, collaterals may also not always secure loan repayment, and loans with collateral sometimes have higher default rates (Jiménez & Saurina, 2004). All this implies a flaw in the solely hard information based credit analyzing model. Moreover, most of the empirical research has suggested that the combination of hard and soft information can achieve a better predicting power of defaults. (Grunert *et al.*, 2005; Godbillon-Camus & Godlewski, 2005; Dorfleitner *et al.*, 2016). Deriving substitution predictive factors for credit analysis becomes significant.⁵ This is the empirical literature to which we shall now turn.

2.2 Soft and hard information in credit analysis- SME empirical literature

Most of the literature concerning soft information has so far dealt with SME financing. The initial studies conducted to explore the role of soft information in credit analysis focused on relationship lending⁶. Existing research has been conducted both on the empirical as well as

⁵ For those reasons, small loans that are characterized as micro loans have been increasingly based on a "high touch" methodology known as "character" lending. The methodology is based on the assumption that the associated borrowers do not have financial statements and may not have collateral. The credit appraisals for these microloans are seriously constrained by "low-tech" processes.

⁶ In the early days of microfinance, virtually all lending was based on group lending, solidarity groups in Grameen's terminology. That is, clients of MFI would form social groups of borrowers, and loans would be attributed individually to each member of the grouping with an identical amount. If the borrower defaulted, the entire group was considered in default which leads to cross default provisions in the loan terms. Group lending leads to social solidarity because the groups assist members who are ill or had other serious difficulties to service their loans. This is an advantage of group lending.

the theoretical level. In the empirical literature pertaining to microfinance, different researchers have utilized different proxies to measure relationship lending; a few have used the distance between the lender and borrower (Dell’Ariccia & Marquez, 2004; Berger *et al.*,2005; DeYoung *et al.*,2008 ; Agarwal & Hauswald,2010), whereas a few others used small bank loans based on soft information as the proxy for relationship lending(e.g., Berger & Undell, 2002). The theoretical work tends to draw on the principal-agent framework to explore relationship lending (e.g., Godbillon-Camus & Godlewski, 2005).

The use of “relationship” indicators for lending purposes have been criticized on several grounds. First, since the loan agent is the repository of the soft information, there is an “agency” problem between the loan agent and the bank management (Berger &Undell,2002). Second, researchers have chosen different proxies to measure the “relationship”. Third, due to the lack of real data, empirical work has remained relatively limited.

Berger and Undell (2002) investigate relationship lending under different bank structures. They divide the small business lending methods into four kinds: lending based on assets, financial statement, credit scoring and the relationship lending. The first three are grouped into transaction-based lending which relies on “hard” information that is easily accessible before the loan is granted. However, relationship lending is based on the lender’s relationship with the borrower, which is relatively “soft.” This soft information is collected through long-term contacts with small enterprises and their local communities. Such information is valuable, but it is time-consuming to collect, and it is hard to convey. The key features of relationship lending have been identified: the loan agent is the repository of the soft information, and an agency problem exists between the loan agent and the bank management. In addition, their results suggest that small banks suit the relationship lending better due to fewer management layers. Similar results were obtained by Uchida (2011) who used Japanese SME data to investigate the factors that influence commercial bank loans. His findings show that smaller banks attach more importance to relationship and collateral factors even though banks under pressure also rely more on relationship factors.

Godbillon-Camus & Godlewski (2005) examine the impact of soft and hard information on credit risk management in a principal-agent framework in which bankers manage credit risk through equity allocation for value at risk coverage. They derive the optimal level of a given banker’s expected utility under both the hard information scenario and the combination of hard and soft information scenario. They show that the combination of hard and soft information reduced the value at risk and thus realized the optimal solution.

Agarwal et al. (2011) analyze the role of soft information in home equity loans based on a dataset containing 108,117 home equity applications in the US. A dummy variable, whether

the applicant was offered a different APR as the counter offer, is used as the proxy for the soft information for the reason that the lower APR is usually due to the secondary screening by the relationship lending. The results show that loans with additional screening and lower APR have 11.1% less default probability. This suggests that the use of soft information in credit analysis for home equity loan effectively reduced the overall portfolio credit losses.

Garcia-Appendini (2007) defines soft information as any kind of data other than transparent public information. Length of the relationship maintained with the borrower is used as the proxy for the relationship lending, and the results show that private information gathered from the relationship plays an important role in the bank lending decisions. If there is no private information from the previous relationship, then the reliance on the firm's reputation is increasing. Reputation can be regarded as a kind of soft information as well in our research. Grunert et al. (2005) use credit file data from 4 major German banks to investigate the role of non-financial factors in predicting default events. The results show that the combination of the financial and the non-financial factors can have better-predicting results.

With the development of the big data and cloud computing, the collection and analysis of soft information has become possible and grow in importance. For example, Xu and Zou (2010) investigate the process of bank lending to SMEs and the incentive mechanisms in China based on the hard and soft information. Using the data of Chinese big four State-owned commercial banks in 12 provinces in China with 239 branches; they build a principal-agent model to empirically test the asymmetric information problem between banks and the SMEs. Their study shows that despite the availability and transferability of both hard and soft information, it is only hard information that is conveyed to the headquarters. Soft information was not shared by branches. The authors suggest, therefore, that branches should be given the right of releasing loans to SMEs, and incentives to avoid moral hazard. Their empirical evidence also suggests that the "harder" the information (and, therefore, the greater transparency of the financial statement of SMEs), the less decision power will be given to branches. This research seems to suggest that soft and hard information have different signaling properties and can influence the division of the decision power between headquarters and branches towards loans to SMEs.

In contrast, reliance on hard information has not generated the results which would close the debate about the quality of information available in lending to SMEs. For example, Berger *et al.* (2005) investigate the relationship between the SBSCS (Small Business Credit Scoring) and the quantity, price, and risk of small business loans based on survey data of U.S large banks. The results show that the credit scoring system has a significant impact on the lending behavior of the banks that adopted the technology. These banks tend to make more small business loans but the average price is higher and the risk level of the loans is higher as well.

These results suggest that the credit scoring system contributes to increasing the credit availability to small business, but its function for credit risk analysis raises questions for the efficiency of the system. A similar result has been obtained by DeYoung *et al.* (2008), who found a positive relationship between the use of credit scoring model and the loan defaults rates. The author suggests that this indicates “credit scoring underwriting process collect less information and thus lead to more loan approval errors.” Furthermore, collateral, which is widely used as one of the most desirable factors in credit approval cannot always secure repayment behavior, As shown, for example, by Jiménez & Saurina (2004), loans with collateral are actually sometimes associated with higher default rates.

To sum up, the SME literature concerning soft and hard information in empirical credit analysis shows that the credit scoring system in SME lending is not foolproof and does not prevent defaults or even mispricing of credit risk. Most of the empirical research has suggested that the combination of hard and soft information can achieve better predicting power of defaults. (Grunert *et al.*,2005; Godbillon-Camus & Godlewski,2005; Dorfleitner *et al.*,2016). While the credit scoring system could undoubtedly always be improved, the SME literature points to another direction towards improving credit appraisal - the combination of hard and soft information may achieve better results in predicting default.

2.3 Hard and soft information in the P2P lending

With relatively open and transparent data from the P2P lending market, the latter opens a possibility to analyze the creditworthiness of the borrower and the investment behavior of the lender. Klafft (2008) use 54077 listings on the Prosper platform and investigate the determinants of the success of a loan and the interest rate. The results show that the verified bank account information and the credit rating are the key determinants. Iyer *et al.* (2009) explore whether the lenders can infer the creditworthiness of borrowers from the aggregate credit category instead of an exact credit score. The results show that the lender can differentiate the creditworthiness of the borrowers with different credit scores but only in the same credit categories. The lenders make their judgment mostly based on “hard” information, such as debt to income ratio, but the research shows that while hard credit information is useful, its usefulness is relatively limited. Serrano-Cinca *et al.* (2015) try to find out the key determinants of the probability of default for P2P lending, and the results show that except for the loan features (interest rate, amount, term), the probability of default is also significantly related to the applicant’s annual income, housing situation, credit record, and indebtedness.

The role of soft information in credit appraisal and loan defaults in P2P markets remains

relatively unexplored⁷. The original ideas go back to early 1990's when psychologists together with economists explored the importance “softer” psychological factors in the analysis of consumer finance. For example, Lea *et al.* (1995) examined the effects on credit appraisal of factors such as personal feelings about the borrowers' financial status, their family support or lifestyle to match with their real debt level. Their findings show that the psychological factors are significantly related to the debt level. Later studies, such as Bertrand *et al.* (2005), look at variables such as gender and race to analyze their effect on the consumer credit appraisal. They find a significant impact of psychological factors in credit risk analysis. For instance, male and female borrowers had 1.3 to 2.2 percentage point difference on the interest rate charged by lenders. They suggest that standard economic models may be missing important psychological drivers of lending. A deeper understanding of specific drivers of lending is vital to improve the lending performance in the P2P and other microfinance markets.

The second important stream of literature dealing with soft information in the P2P markets is lending based on social capital. The central proposition of social capital theory is that an individual's social relationships have social values in conducting the social affairs (Bourdieu 1986). Greiner and Wang (2009), for example, examine the theory by investigating the impact of social capital from both the borrower's and the investor's perspective. Members of the P2P lending market have two ways of building social capital: endorsing each other to create networks of friends or join groups. The results show that from the borrower's perspective, social capital has a positive impact on the success rate of the funding and it plays a more important role in groups with low credit rating. From the lender's perspective, social capital has a positive influence on the loan payment performance, but the impact is quite small.

Another example of a study concerning social networks lending is Everett (2011). He used the probit model based on prosper.com data to examine the impact of social connections on default rate and interest rate. His results show that membership in a group with private information is a sign of lower default possibility. Liu *et al.* (2015) using data from Chinese P2P lending platform PPDai.com evaluates the impact of friendship on the success rate of P2P lending. The paper defined friendship from three aspects: the pipes, the prisms, and the relational herding. “Pipes” refer to the friends of the borrower who were divided into offline friends and online friends. The results show that an offline friend is more likely to bid on

⁷ The use of soft information is not unknown in the other arm of Fintech industry – in non-bank financial institutions. Those institutions rely on proprietary models and use a combination of hard and soft information to evaluate credit risk. Their activities have increased dramatically but they remain relatively high-cost since they are paying commission agents to bring potential clients to them.

listing than online friends. This may be due to the closer relationship between offline friends. “Prisms” refer to situations in which the borrower may or may not have a bid from a friend. The results show that a bid by a friend has a negative impact on investor’s judgment of investing in the applicant. The reason is that the lender believes that the bid of the borrower usually contains private information which is only offered to a friend. Relational herding refers to the phenomenon in which the lender is more likely to follow the bid that the lender’s friends have invested in. However, the empirical tests of these hypotheses are disappointing. Nevertheless, a further breakdown of different types of friendship showed more promising results - those lenders were more likely to follow offline friends than online friends and strangers. This would suggest the existence of a “herding” behavior.

Cao (2013) using the Chinese first P2P platform PPDai data analyzed the impact of social capital on the success and default rate of the lending outcome in P2P industry. He chooses the degree of centralization level of the social networks, a method from SNA (social network analysis), as the proxy for social capital. The results show that the borrower with higher social capital will have higher chances to be funded and with lower default rate. From the lender’s perspective, people with higher social capital will be more active in bidding and listing, and will also have better chances of successful bidding.

An interesting aspect of lending and the role of soft information in P2P markets is provided by Herrero-Lopez (2009). He argues with the help of empirical evidence that social capital positively impacted the success rate of the funding when the hard information was not sufficient, but noted that social networks wouldn’t eliminate the default risk. Freedman & Jin (2011) argue that endorsements and bids from friends indicate lower default rate and higher returns. This may due to the private information borrowers shared with their friends can which, in turn, help them to analyze the risk and make more appropriate decisions. They also proposed that monitoring within social networks will reduce the delay of the payment thus reduce the default chances.

Finally, Miu & Chen (2014) examine trust as a factor in the form “recommendations” - “recommended” trust (friends provide reference letter) as the proxy for social capital and thus break the limitation of using group relation and friends’ numbers as the social capital variables. The results show that borrowers with higher social capital have lower default rates due to the supervision from the social circle and the pressure of losing trust. Similar results are obtained by Lin *et al.* (2013). Using also Prosper.com dataset, they show that friendship has a significant impact on the funding success rate and ex-post default rate. Friendship had a positive relationship with the funding possibility and negative relationship with the default probability.

Perhaps one of the most frequently analyzed aspects of soft information is the impact of appearance and race in credit analysis. Given the size of the literature, only the most representative studies are included in this brief review. Using the data of Prosper.com in 2007 between March and June with 2012 borrowers and 18627 lenders, Ravina (2012) analyzed the impact of beauty and personal characteristics in credit markets. When controlling for hard information like credit history and financial situation, the results show that borrowers' appearance and personal characteristics had a significant impact on the success rate of the loans and the interest rate. People with beautiful profile pictures had 1.59% higher chances of getting a loan and with 60bps lower interest rate although they have similar default rate as the average looking people. Black people were less likely to get a loan. Blacks, Asians, and older people were paying a higher interest rate.

Dorfleitner *et al.* (2016) used data from two German P2P lending platform Smava and Auxmoney to investigate the power of text related soft information in P2P credit analysis. They tested the problem in three aspects; the typos, the length and the keywords. The results showed that text related factors did influence the decision of investors, but had no significant relation with the default probability.

Gonzalez & Loureiro (2014) investigated the impact of the borrower's profile pictures on the probability of getting a loan. The results also show that the perceived age is a signal for working experience which, in turn, results in a higher loan amount granted. Needless to point out that the result goes against Ravina's "beauty premium" theory and they support the "beauty is beastly" effect. The results suggest that the attractiveness of the profile picture has no positive influence on the loan success; it actually may have a negative impact when the lender is of the same gender.

Duarte *et al.* (2012) also look at the issue of appearance on the P2P lending market. They used a psychometric method to define the trustworthy appearance of the applicants. The results show that people who look more trustworthy will have higher chances of getting their loan funded and that they will also have less default rate. Pope and Sydnor (2011) separate the discrimination in the P2P lending industry into taste-based and statistical based. Taste-based discrimination is purely based on lender's favor, and the statistical discrimination present the discrimination based on the borrower's default possibilities such those of old people. They found that the successfully founded African American loans were 25% to 34% smaller than those of whites. Although the interest rate of African American loans is 0.6 to 0.8 higher than the white loans, the net return on the African American loans is significantly worse than those on the white people's loans. Other soft information also included groups and friends or even the profile pictures. Once again, the results confirmed the contention that borrowers without profile picture or with unhappy-looking profile pictures were less likely to get a loan.

Liao *et al.* (2015) attempted to investigate the influence of diploma (education) on the creditworthiness of borrowers in Chinese P2P platform RenrenDai. His results show that borrowers with high educational level are less likely to default. However, the investors showed no preference for these borrowers which implies that borrowers with higher educational attainments in China did not get the access to credit easier.

Finally, Barasinska (2010) used data from German's largest P2P lending platform Smava.de to analyze gender as a key determinant for the success of a loan. The results show that gender does not have a significant impact on the loan success which is in contrast to the results of Ravina (2012) and Pope&Sydnor (2011). This strongly suggests that the P2P markets in different countries are heterogeneous.

A Brief Summary: Studies of credit appraisal on P2P platforms provide some evidence about factors which determine the success of loan application as well as the loan performance in the market. As in other segments of financial markets, hard information and credit scoring in P2P lending play an important role in the creditworthiness analysis and the success of the loan. Financial factors like income, housing situation, indebtedness have been shown to have a significant relationship with the likelihood of success in getting a loan and the probability of default. However, hard information is often not available in these markets, so other approaches need to be taken. Moreover, defaults cannot be avoided even with standard hard information.

As in the SME literature, the alternative and complementary directions of credit appraisals in the P2P markets have been increasingly focused on the use of soft information. The vast majority of the existing literature concerning the soft information is about the role of social capital, and most of the research shows that social capital has a significant positive impact on the funding success rate. However, the results are far less clear about the value of soft information in predicting the borrower's repayment performance. Some studies show that friendship online is a sign of lower probability of default. But other studies find that social networks cannot secure the repayment of loans. Similarly, contradictory results have been reached with regards to the role of appearance, language, and gender in the success of the loan application as well as in predicting defaults. Perhaps the most reliable variable predicting the loan performance has been shown to be race and age. It is, therefore evident that the empirical evidence about the role of soft information in credit appraisal in the P2P market has so far been very mixed and inconclusive.

The inconclusive nature of the studies is partly due to the fact that most studies primarily deal with the question about effect of a particular soft indicator such as gender, race, age or social

networks on the success of loan applications or the terms of loans. s) and on the effects of a single soft variable (Miu and Chen 2014 , Ravina 2012, Barasinska 2010, Serrano-Cinca *et al.* 2015, Potzch and Bohme 2010). Studies of defaults primarily focused on the effects of soft information on debt level (e.g. Lea *et al.* 1995). No systematic attempt has been made to study the role of a wider set of “soft” determinants of defaults Methodologies and approaches of studying the role of “soft” information have been different being driven by different questions of researchers. In addition, empirical studies covered different markets, and one conclusion that can be made from the empirical literature is that the P2P markets are not homogeneous and are most likely made up by different institutional structures, cultures and psychological factors. Finally, some of the contradictory results can be explained by the reliance on a very narrow range of explanatory variables in determining defaults.

3 Methodology

A binary classification model is used to assess the value of soft information in credit appraisal is using a case study approach. We shall build a model which will be tested by drawing on the Chinese P2P platform and data available from the platform. We shall start with a brief description of data, followed by a description of the model. Since the model will be tested in different variants, the description will also include an explanation of our analytical treatment of model discrimination.

3.1 Approach, Data, Scope and Definitions

As noted in the previous section, loan determinants have been typically studied in the context of loan applications and loan terms. Our approach is different; we shall examine the role of determinants from the perspective of loan defaults. This primarily reflects the explosion of P2P lending in China and with it the credit risk and the likelihood of defaults.⁸ Defaults in microfinance have been studied primarily in the context of hard information. Our approach is to focus on the role of soft information and in combination with the role of hard information. In addition, we shall examine a combined effect of different soft variables, in contrast to studies which only focus on a single soft variable.

⁸ The emphasis on loan appraisal could be justified in the past by the relatively successful performance of microfinance lending. According to recent estimates, portfolio at risk 30 and 90 days and loan lost rates have remained very low in the microfinance industry, with losses at 1 percent or lower throughout the world, with some variations by regions. However, after the explosion of P2P lending, the credit risk is rising. For more details, see Lieberman *et al.* (2017).

As we have also seen in the previous section, definitions of soft information differ. We shall define soft information as information transmitted by a selected social or psychological “characteristic” of borrowers. The information consists of two parts – information identifying the characteristic of the borrower and information identifying the intensity of the effect of that “characteristic” on defaults. While the intensity of the effect of the given “characteristic” on loan defaults is difficult to quantify, the characteristic of the borrower (agent) transmitting the information can be identified and verified and is, therefore, used as a proxy. Soft information contains information about the borrowers such as age, education, gender or race and also information based on judgments about the personality and reliability of the borrower, the state of the business seeking a loan, the potential for future development, etc. Soft information may even include a borrower’s social networks, video interviews, profile picture, description of the borrowing stories, the definition allows to capture a link between the relevant characteristic of the agent and defaults. This is a valuable property even though it does not allow to capture the intensity of the relationship. Our definition is consistent with the manner in which it has been defined in the literature (e.g. Grunert, J., Norden, L., & Weber, M. 2005).

The specific choice of soft variables was driven by the availability of data and theory and guided by the literature. Our treatment of soft information includes social indicators such as duration of education (Liao *et al.* 2015), age (e.g., Gonzalez & Loureiro 2014) and gender (e.g., Barasinska, 2010, Ravina, 2012 and Pope & Sydnor, 2011). Those are indicators that can be identified and verified. We have also added other types of soft information such as variables which refer to personal characteristics and social networks of borrowers and those, in turn, represent other proxies for social capital and networks. Due to the limitations of data we were unable to use other soft indicators but believe that we have been able to capture a sufficiently broad range of those variables which have been most frequently used in the literature.⁹

We shall examine the role of soft information using a case study of the P2P market in China, and this is done for two reasons: (1) The paper uses the data from the Chinese P2P platform RenrenDai.com. The case of the Chinese P2P market is compelling because of its size and rapid growth. Moreover, the market has developed hand-in-hand- with the development of a rich database which includes a precious source of soft information.

Moreover, the Chinese market is interesting because of its institutional specifics. The system provides for a verification of mobile telephone users which allows tracing and verifying the real users of cellphones. This, in turn, increases transparency of borrowers and helps enhance trusts in the information provided by borrowers. In addition, like many other emerging markets, China has a short investment history and relatively low public financial literacy and credit analysis based on a broad range of indicators becomes of utmost importance in this market. However, China has a powerful oversight mechanisms in the financial sector which

⁹ For example, one could use as proxies for soft information “geographic distance” between the lender and the borrower or the “length of the relationship” between the lender and borrower but those data are, alas, not available on the platform.

operates both on the formal and informal level.¹⁰ Furthermore, the contractual arrangements in China are heavily influenced by Chinese culture which favors information derived from human relationships. Soft information has, therefore, become a special requirement for contracts and for P2P markets in China in particular.

The dataset used in this paper is from the Chinese Platform RenrenDai. It is a Chinese P2P lending platform which is established in 2010. The choice of China was made for two reasons. First, the case of the Chinese P2P market is compelling because of its size and rapid growth. According to National Bureau of Statistics of PRC, the transaction volume of P2P markets has been reaching 2.8 trillion yuan at the end of 2017 and the market contains 1931 platforms¹¹. Second, the Chinese market has developed hand-in-hand- with the development of a rich database which includes a precious source of soft information. The platform is one of the pioneer Fintech companies. Until October 2016, the total amount of transaction has exceeded 21.2 billion yuan. The platform is targeting at microloans; the average loan amount is 71,000 yuan. The platform consisted of 251,887 listings from 2010 to 2014. Since we are analyzing the predicting power of two different kinds of credit information, the failing auctions, the listings that are still in bidding process or have not yet been repaid have been eliminated. Only the 14,575 finished listings have been retained. The summary of the listings is shown in Table 3.1. The description of dataset concerning hard and soft information is provided further below.

Table 3.1 Distribution of the listings

Over Due	84
Wait for Open	11
Default	590
Failing Auctions	18,1043
Finished the repayment	13,901
Under application process	5,439
Under repayment process	50,819
Total	251,887

¹⁰ The oversight of Chinese banks continues to be closely linked not only with government but also with the political structures. It is also well known that misinformation or corruption can have dire consequences for malfasants in the Chinese markets. See, for example, Jianying Zha: *Tide Players: The Movers and Shakers of a Rising China*, New York, New Press, 2011.

¹¹ For more information, see Annex 1.

3.2 Model

Our model starts from the premise of a key role played by hard information and the model includes, therefore, a range of standard hard information. However, following the work on hidden information under imperfect contracts (Stiglitz,1981; Spence,1973; and Akerlof, 1970), soft information variables can be seen as an additional element of information about borrowers in addressing the “asymmetric information. “problem. In addition, we expect a better loan performance with soft information because the latter helps build trust between lenders and borrowers. As Potzch and Bohme (2010) show, trust can even lead to better credit conditions and non-financial personal data can improve the trust situation.¹² More recently, Thabor and Mentor (2018) analyzed the competitive interactions between banks and non-bank lenders and distinguishing between trust and reputation they showed that trust enables lenders in (Fintech firms) to have assured access to financing, whereas a loss of investor trust makes the access conditional on market conditions and lender reputation. They further show that banks have stronger incentives to maintain trust. When borrower defaults erode trust in lenders, banks are able to survive the erosion of trust when Fintech lenders do not.¹³ More corroborative evidence on the importance of trust enhanced by soft information has been provided by Miu and Chen (2014) , Ravina (2012), Barasinska (2010), Serrano-Cinca et al. (2015).

In order to assess the role of soft information in credit performance I test a model in which the determinants of defaults are hard and soft information in different scenarios. In the first scenario, we shall only consider hard information as the only set of factors determining defaults. In the second scenario, we shall reverse the procedure and only consider soft information as being the sole determinant of defaults. In the third scenario, we test the joint influence of both hard and soft information. We shall test two hypotheses:

Hypothesis 1: Credit appraisal based on appropriately selected soft information can have a strong predicting power;

Hypothesis 2: The credit predicting model can be strengthened by soft information. Soft information can capture useful information for credit analysis that has been neglected by the financial hard information.

¹² “Our results indicate that communicating personal data indeed supports a positive and trustworthy overall perception of the borrower, but the impact on credit conditions is marginal: only data in the category education, profession, and qualifications show a small and significant effect.”

¹³ In their theory of trust in lending they make a crucial point about a difference between banks and Fintech lenders. Banks have access to insured deposits and provide valuable depository services to their customers whereas Fintech platforms are typically investor-financed. This means that the distinction makes banks more trustworthy than Fintech platforms, gives banks a competitive advantage and that the Fintech lenders are critically dependent on maintaining trust.

In order to estimate the probability of default, a binary regression estimation model-Logit Regression is chosen. Receiver operating characteristic (ROC) curve will be used to compare the performance of soft and hard information models as one way of discriminating among different estimates.

Model 1: $Y_i = \alpha \text{ Hard Information} + \alpha \text{ Control Variables} + \varepsilon \dots\dots\dots(1)$

Model 2: $Y_i = \alpha \text{ Soft Information} + \alpha \text{ Control Variables} + \varepsilon \dots\dots\dots(2)$

Model 3: $Y_i = \alpha \text{ Hard Information} + \beta \text{ Soft Information} + \alpha \text{ Control Variables} + \varepsilon \dots\dots(3)$

Y is the dependent variable which represent whether the loan has been paid back completely without delay. “1” represents “default”; “0” represents “repaid”. The control variables include the interest rate, the length of the loan, and the amount of the loan.

Table 3.2 Description of independent variables

Variables	Description
Hard Information	
Income Level	Category variable: Monthly Income of the borrower (1~5) Group 1: <1000 yuan Group 2: 1001~2000 yuan Group 3: 2000~5000 yuan Group 4: 5000~10000 yuan Group 5: 10000~20000 yuan Group 6: 20000~50000 yuan Group 7: >50000 yuan
Income verification	Dummy variable: the income is verified -1; is not verified-0
House ownership verification	Dummy variable: the ownership is verified-1;is not verified-0
Car ownership	Dummy variable: the borrower has a car-1;doesn't have-0
Car ownership verification	Dummy variable: the ownership is verified-1;is not verified-0
Mortgage Loans	Dummy variable: the borrower has a mortgage loan-1;doesn't have a mortgage loan-0
Soft Information	
Loan description	Length of the loan description
Age	The age of the borrower
Gender	Dummy variable: female-1; male-0
Marriage status	Dummy variable: married-1; otherwise-0
Educational Level	Years of taking education
Weibo verification	Dummy variable: the social network is verified-1; is not verified-0
Mobile verification	Dummy variable: the mobile number is verified-1; is not verified-0
Video verification	Dummy variable: finished the video verification-1;otherwise-0

Loan Features	
Interest	The interest rate of the loan
Term	The length of the loan
Amount	The amount of the loan

The proxies for the hard information in our model are the key financial determinants which indicate the wealth and solvency of the borrower. The financial determinants we choose are the three key fundamental financial indicators that are available in our dataset: the monthly income, the house ownership, the car ownership, and the mortgage loans. The car and house ownership are two dummy variables with the value of “1” for “ownership” and “0” for “doesn’t have”. We use the interaction terms of car ownership and verification of the ownership to certify the accuracy of the data. Following Order & Zorn (2000), who found that default and losses were higher in low income groups, we have also chosen monthly income as one of the independent variables. We also put the verification of the income into the model to certify the accuracy.

As soft information is difficult to measure, it is necessary to use proxies. The proxies used in our model are summarized in Table 3.2. Due to the limitation of the dataset, we cannot obtain data concerning the discussion groups on RenrenDai.com platform. Thus, the verification of Chinese largest social network weibo is chosen as the second best option and as the indicator for the social impact. According to Weibo 2016 Development Report, the active users of Weibo were 297 million at the end of September 2016. This guarantees the choice of Weibo verification as a social image proxy. If the applicant’s social network was verified then it is represented as “1”, otherwise “0”.

The Chinese P2P lending platform does not usually provide real pictures as the profile picture of persons. We have, therefore, chosen the video verification as the proxy for the image indicator. At the same time, it can also be regarded as a social indicator. If the applicant agreed on a video verification, then it is represented as “1”, otherwise “0”. The explosion of mobile services is the soil for the Fintech 2.0, and the mobile usage data is the battle field for Fintech companies, especially for the big data Fintech companies. It is the essential source for the anti-fraud measures since the mobile number has been added to the real-name system in China. This allows tracing and verifying the real users of cellphones. In addition, mobile using behavior has been recognized as one of the most effective indicators of default. Thus we add the mobile verification variable to our model. It is also a dummy variable, “1” equals verified, otherwise “0”.

Following Lin *et al.* (2016), we have also incorporated the impact of language on the bidding and default behavior. We use the length of the description of loan purpose as the linguistic

indicator. The treatment of other soft information is similar as is common in the literature: duration of education (Liao *et al.* 2015), age (e.g., Gonzalez & Loureiro 2014) and gender (e.g., Barasinska, 2010, Ravina, 2012 and Pope&Sydnor, 2011))

3.3 Receiver Operating Characteristics (ROC) Curves

Since our model is estimated in three different versions, we need to determine whether the model estimates can be discriminated among them purely on econometric as opposed to theoretical grounds. A receiver operating characteristic graph is a technique for visualizing and selecting classifiers based on their performance (Fawcett, 2006).

Table 3.3 Four Cases for Binary Classification

		Predicted Class	
		Class1	Class0
Actual Class	Class 1	True Positives	False Negatives
	Class 0	False Positives	True Negatives

As shown in Table 3.3, there are four cases for the binary classification model:

True Positives: The predicted class is 1, and the actual class is 1;

True Negatives: The predicted class is 0, and the actual class is 0;

False Positives: The predicted class is 1, and the actual class is 0;

False Negatives: The predicted class is 0, and the actual class is 1.

The ROC curve is the graphical plot which shows the performance of a binary classifier by diagrammatizing the true positive rate (TPR) against the false positive rate (FPR) at different thresholds. The TPR and FPR are known as sensitivity and specificity classification function in statistics which represent the proportion of positives and negatives of the detection accordingly. The formula for TPR and FPR is as below:

$$TPR = TP / (TP + FN) \dots\dots\dots(4)$$

where TP stands for “true positive” and FN stands for “false negative.” Equation (4) represents the rate of correctly diagnosed number among all positive numbers in the sample. Similarly,

$$FPR = FP / (FP + TN) \dots\dots\dots(5)$$

where FP stands for “false positive” and TN for “true negatives.” Equation (5) represents the rate of wrongly diagnosed numbers among all negative numbers in the sample.

The ROC curve can be plotted by the TPR and FPR ratios against their different thresholds. TPR data are plotted on the vertical axis (sensitivity) and FPR (Specificity) data on the horizontal axis. An important parameter of the ROC curves is the AUC- the area under the curve. AUC as a measure of the accuracy of the classifier, and it represents the probability of the classifier ranking a randomly chosen positive instance higher than a randomly chosen negative instance (Fawcett, 2006). The closer the ROC curve is to the upper left-hand or the closer the AUC is to the value of 1, the truer positives are defined, creating a better classifier. The area under the ROC curve is derived as:

$$ROC(AUC) = \int_1^0 TPR(x) FPR'(x) dx \dots\dots\dots(6)$$

4 Results

Our results are presented for the three versions of our model. The predicting power of the hard information on default will be tested at first. We shall then compare the results with those in model version 2 utilizing solely soft information as the key determinant. Finally, we shall see how much the predicting power can improve if we combine both kinds of information – hard and soft information. The logit regression results are presented in the following section, and a comparison of the ROC curves for the three models is discussed in section 4.2.

4.1 The logit Regression Results

Model 1 investigates the relationship between the probability of default and the traditional hard financial indicators. The results are reported in Table 4.1.

Table 4.1 Logit Regression Results for Model 1

VARIABLES	(1) default	(1) default	(1) default
1.Income verified	0.765*** (0.210)	0.775*** (0.219)	-0.263 (0.226)
1.Income	-0.795 (1.015)	-0.629 (1.021)	-0.739 (1.043)
2.Income	-0.0458 (0.310)	-0.905*** (0.332)	-0.493 (0.343)
3.Income	-0.320** (0.129)	-0.355*** (0.131)	-0.360*** (0.135)
5.Income	-0.256 (0.161)	-0.265 (0.166)	-0.360** (0.173)
6.Income	0.370*** (0.132)	0.431*** (0.136)	0.354** (0.139)
7.Income	0.444*** (0.125)	0.523*** (0.133)	0.382*** (0.138)
1.Incomeverified#1.Income	0 (0)	0 (0)	0 (0)
1.Incomeverified#2.Income	1.311 (1.211)	2.341** (1.104)	2.384*** (0.879)
1.Incomeverified#3.Income	0.471 (0.295)	0.487 (0.310)	0.513 (0.320)
1.Incomeverified#5.Income	-1.117** (0.561)	-1.256** (0.569)	-1.178** (0.555)
1.Incomeverified#6.Income	-1.879*** (0.558)	-1.766*** (0.566)	-1.515*** (0.574)
1.Incomeverified#7.Income	-2.518*** (0.557)	-2.342*** (0.563)	-1.913*** (0.578)
Car verified	-0.0832 (0.112)	-0.201* (0.109)	-0.0941 (0.118)
1.House verified	0.601*** (0.124)	0.491*** (0.119)	0.627*** (0.126)
1.Mortgage loan	-0.482** (0.208)	-0.394* (0.218)	-0.525** (0.225)
1.Houseverified#1.Mortgag	-0.378	-0.523*	-0.384

eloan			
	(0.267)	(0.280)	(0.290)
Interest		0.216***	0.274***
		(0.0118)	(0.0139)
Term		-0.0168***	-0.0403***
		(0.00456)	(0.00516)
Amount		-4.39e-07	-1.91e-07
		(4.10e-07)	(3.79e-07)
2011.year			0.417
			(0.726)
2012.year			1.248*
			(0.724)
2013.year			1.876***
			(0.725)
2014.year			3.187***
			(0.734)
Constant	-3.129***	-5.895***	-7.929***
	(0.0975)	(0.221)	(0.772)
Pseudo R2	0.0294	0.0852	0.1226
Observations	14,569	14,569	14,569

Heteroscedasticity-Robust, standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note. The numbers associated with the variable “income” refer to income groups. The sample included 7 income groups.

Table 4.1 presents our logit regression results for Model 1. The model investigates the relationship between the traditional hard credit information and the default behavior. The interest rate, amount and term have been chosen as the control variables to control the omitted variable bias. Since we are using panel dataset, year dummy variables are added to control for heterogeneity in the adjusted model (last column).

Table 4.2 Income Distribution

Group	Income Level	Freq.	Percent
1	< ¥ 1000	51	0.35
2	¥ 1001-2000	312	2.14
3	¥ 2000-5000	4,464	30.6
4	¥ 5000-10000	3,235	22.20
5	¥ 10000-20000	2,013	13.82
6	¥ 20000-50000	2,116	14.52
7	> ¥ 50000	2,378	16.32
	Total	14,569	100.00

Variable income represents the monthly income of the borrowers; the seven categories of the income are shown in Table 4.2. The median income group (¥ 5000~10000) is chosen as the reference group for category variable income. The interaction effect of income and verified income is significant except in group 3 and the omitted group¹. The coefficient proves that borrowers who earn ¥ 1001~2000 are more likely to default than those who are in the reference group. Borrowers who have more than ¥ 10000 monthly incomes are less likely to default than the borrowers are in the reference group. Car ownership – as an indicator for higher financial status - is insignificant in the model so it doesn't necessarily need to be regarded as a significant indicator of the default behavior.

Some interesting results come up for the effect of house ownership. House ownership certificate turns out to be significantly positively related to the default behavior. This may indicate that the traditional real estate collateral cannot guarantee the creditworthiness on the online P2P lending platform or there is an adverse selection problem existed in the online lending market. The finding is also consistent with the results obtained by Jiménez & Saurina (2004). Moreover, the mortgage loan variable is significantly and negatively related to the default behavior. In other words, if the applicant is in debt of mortgage loan, he/she is less likely to default on the P2P lending platform. This, in turn, could be interpreted as indicating that borrowers with mortgage loans care more about their credit standing. The violation of the traditional indicator of default also hints at the need for other important information in the internet lending market. The goodness of fit indicator (Pseudo R²) is increasing along with the adding of control variables and year dummies. The same feature is consistent with the log likelihood estimations. In general, the results shows that hard financial factors which represents the wealth and solvency of the borrower cannot predict as good as expected and some even has opposite results as expected in the P2P lending market.

Model Version 2 analyzes the relationship between the probability of default and soft credit information and the results are presented in Table 4.3.

Model 2: $Y_i = \alpha \text{Soft Information} + \alpha \text{Control Variables} + \varepsilon$

Table 4.3 Logit Regression Results for Model 2

VARIABLES	(1) default	(1) default	(1) default
Loan description	-0.00647*** (0.000532)	-0.00641*** (0.000551)	-0.00562*** (0.000546)
Age	0.00174 (0.00558)	-0.000483 (0.00601)	0.00480 (0.00596)
Gender	-0.317** (0.126)	-0.262** (0.128)	-0.231* (0.129)
Marriage	-0.353*** (0.0972)	-0.266*** (0.0999)	-0.202** (0.101)
Educational	-0.122*** (0.0155)	-0.117*** (0.0160)	-0.122*** (0.0165)
Mobile verified	-0.555*** (0.126)	-0.523*** (0.129)	-0.639*** (0.132)
Weibo verified	-0.802*** (0.147)	-0.701*** (0.150)	-0.453*** (0.154)
Video verified	0.908*** (0.113)	0.936*** (0.120)	0.976*** (0.123)
Interest		0.191*** (0.0132)	0.242*** (0.0144)
Amount		0.0687* (0.0400)	0.0609 (0.0444)
Term		0.0102* (0.00536)	-0.00653 (0.00595)
2011.year			0.423 (0.740)
2012.year			0.929 (0.739)
2013.year			1.403* (0.737)
2014.year			2.257*** (0.746)
Constant	0.0863 (0.351)	-3.535*** (0.574)	-5.545*** (0.943)

Pseudo R2	0.1127	0.1483	0.1694
Observations	14,571	14,571	14,571

Heteroscedasticity-Robust, standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4.3 presents the logit results for model version 2 which analyzes the relationship between the soft information and the probability of default. As shown, the length of loan purpose description is negatively related to the probability of default, and the results remain consistent after adding the controlling variables and the fixed effects of the year. This means that the more words the applicant wrote on the loan purpose description, the less likely he/she will default. Results for the effects of gender are consistent with the literature and show that women are less likely to default than men. Marital status and educational level are also significant variables. Since we use the length of education to express the educational level, the results show that the longer period spent by the applicant in training or schooling, the less likely he/she will default on the P2P loans. Marital status is also significant before and after the robustness treatment which illustrates that people who have a family are less likely to default. The three social capital variables under consideration are also all significantly related to the probability of default. Mobile verification and social network verification are negatively related to the probability of default. However, video verification is positively related. The only other variable which turns out insignificant is the age which clearly is not consistent with most findings in the literature and would require further investigation.

Table 4.4 Logit Regression Results for Model 3

VARIABLES	(1) default
1.Income verified	-0.184 (0.231)
1.Income	-1.146 (1.196)
2.Income	-0.268 (0.351)
3.Income	-0.146 (0.137)
5.Income	-0.389** (0.173)
6.Income	0.284* (0.150)
7.Income	0.283* (0.157)
1.Income verified#1.Income	0 (0)
1.Income verified#2.Income	2.764*** (0.803)
1.Income verified#3.Income	0.409 (0.336)
1.Income verified#5.Income	-1.135* (0.583)
1.Income verified#6.Income	-1.548*** (0.594)
1.Income verified#7.Income	-1.891*** (0.578)
Car verified	-0.295** (0.116)
1.House verified	0.455*** (0.128)
1.Mortgage Loan	-0.573** (0.225)
1.Houseverified#1Mortgage loan	-0.0162 (0.287)
Loan description	-0.00537***

	(0.000560)
Age	-0.00171 (0.00623)
Gender	-0.254** (0.129)
Marriage	-0.130 (0.106)
Educational	-0.121*** (0.0171)
Mobile verified	-0.579*** (0.136)
Weibo verified	-0.403** (0.157)
Video verified	1.006*** (0.127)
Interest	0.243*** (0.0151)
Amount	0.00969 (0.0497)
Term	-0.00298 (0.00637)
2011.year	0.343 (0.743)
2012.year	0.831 (0.743)
2013.year	1.386* (0.742)
2014.year	2.522*** (0.754)
Constant	-4.903*** (0.976)
Pseudo R2	0.189
Observations	14,566

Heteroscedasticity-Robust, standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note. The numbers associated with the variable “income” refer to income groups. The sample included seven income groups.

Table 4.4 presents the logit regression results with the combined effect of soft and hard independent variables. The significance and the direction of all variables remain consistent with the previous Model Version 1 and Model Version 2 except for the effect of car ownership which becomes significantly negatively related to default. The pseudo R2 is increasing from the 0.123 (Model1) and 0.169 (Model2) to 0.189 (Model3). This suggests that the combination of the hard and soft information can better predict the loan performance.

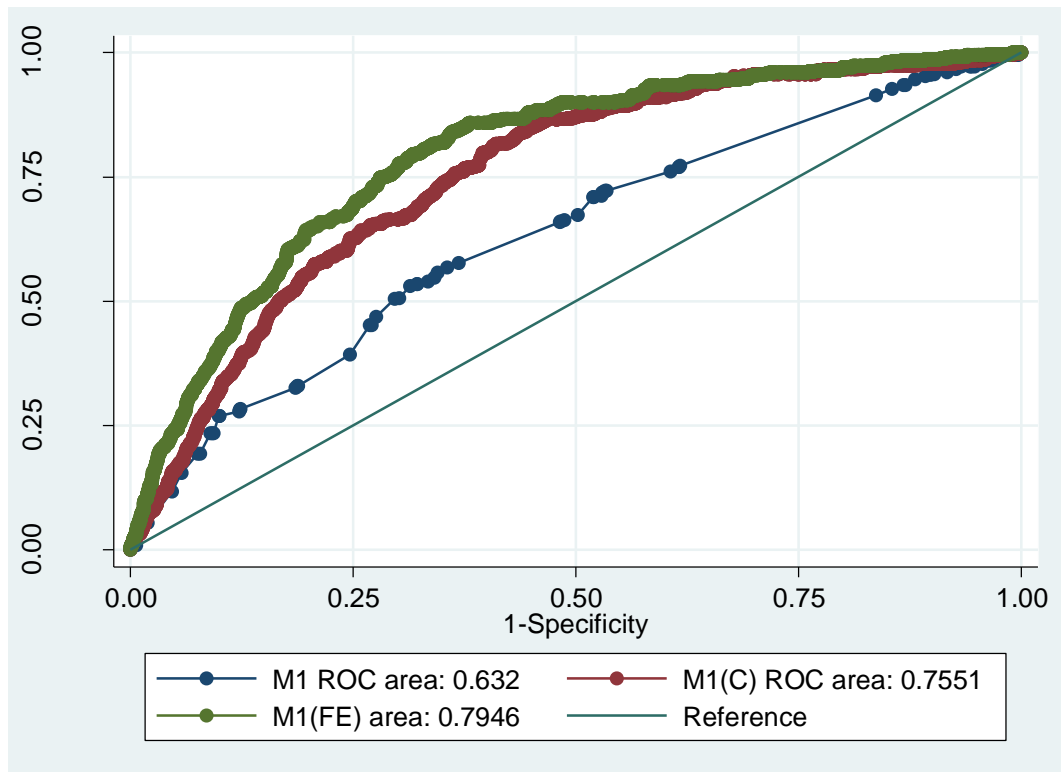
Multi-collinearity could be suspected in this model. We have, therefore, carried out a multicollinearity test, and the results are presented in Appendix 2. The mean variance inflation factor is 1.4. VIF for all variables are less than 2. This suggests the absence of multicollinearity in the model.

4.2 Model Discrimination

All three versions of our model generated significant results for most of the tested variables. Is there a way for us decide which of the models performs best? Before addressing this issue from a theoretical point of view, we turn to the ROC statistical technique described in the previous section. We shall use the ROC curves to measure the performance of the default predicting model. Visually, the more the curve on the graph approaches the upper left-hand corner (0,1), the better the performance of the model. An alternative way of assessing the performance of the estimations is to look at AUC, it is increasing by adding “better” information.

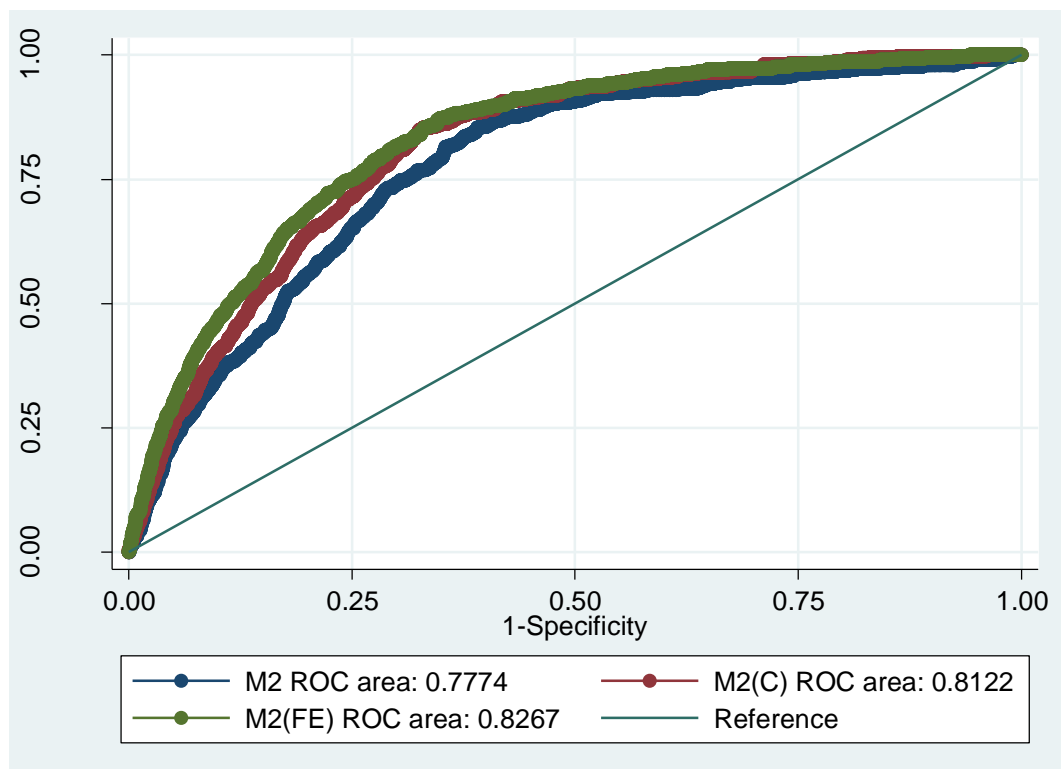
We have generated three ROC graphs corresponding to our three models which are presented in Figures 5.1-5.3. ROCs derived from Model 1 (hard information) is shown in Figure 5.1, ROCs derived from Model 2 is shown in Figure 5.2. ROC in Blue represents the curve from the basic model (hard information, soft information respectively) without control variables and dummy for years. ROC in red represents the basic model plus control variables, and ROC in green represents the basic model plus control variables and dummy variable for years. Figure 5.3 presents the robustness model with control variables and year dummies for Model 1(blue), Model 2(red), Model 3(green).

Figure 5.1. ROC Curves for Model



Starting with Figure 5.1, the AUC in Model 1 is increasing by adding the controlling variables and increasing even more by adding the year dummy. This is also accordance with our results from the Pseudo R2 of Model1.

Figure 5.2 ROC Curves for Model 2



As in Model 1, the AUC in Figure 5.2 is increasing by adding the robustness treatment variables. However, the difference is that the growth interval is not as big as Model1.

Figure 5.3 ROC Curves for Model Comparisons

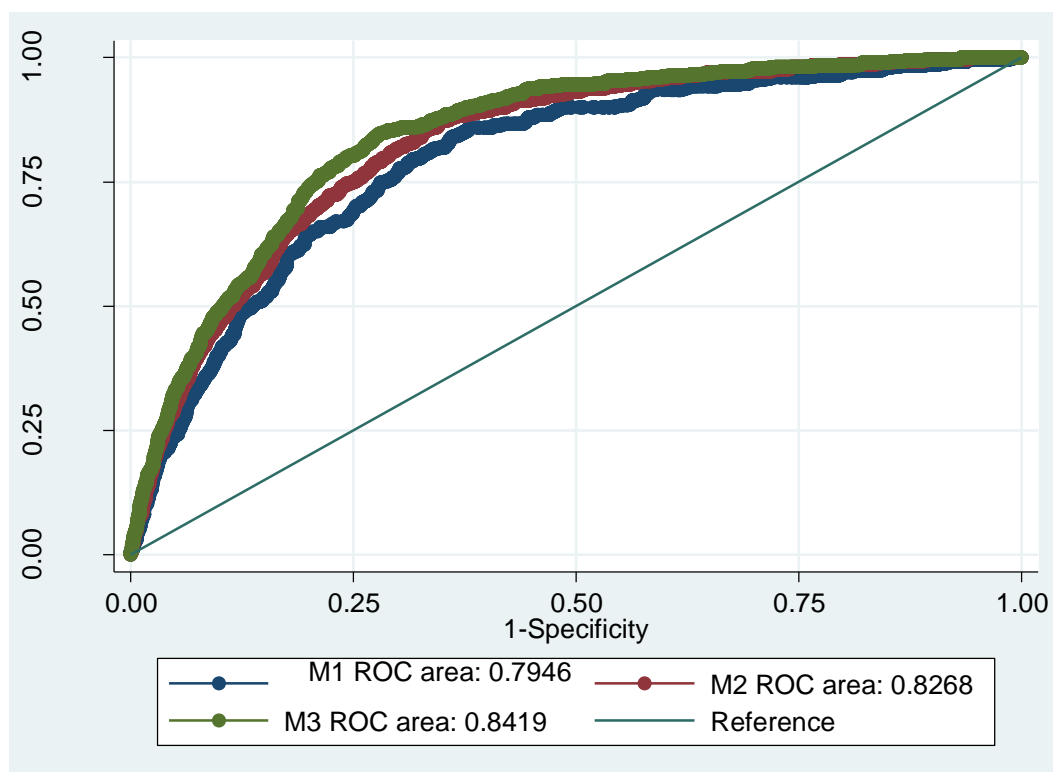


Table 5.1 ROC results of hard and soft information models

Obs	Area	Std. Err.	[95% Conf. Interval]
-----	------	-----------	----------------------

Hard	14,566	0.7946	0.0084	0.77819	0.81098
Soft	14,566	0.8268	0.0073	0.81249	0.84107
Combined	14,566	0.8419	0.0069	0.82829	0.85553

Ho: area(Hard) = area(Soft)=area(Combine)

chi2(1) = 133.48 Prob>chi2 = 0.0000

The computations of AUC are summarized in Table 5.1. Recalling equation (6) above we have calculated and compared AUC in Model 3 (curve related to hard and soft information combined) with that of Model 1 (hard information) and Model 2 (soft information) respectively. ROC in Model 3 has the largest AUC; it is 0.0473 larger in comparison to AUC in Model 1 and 0.0151 in Model 2 (Soft). Other interesting results are obtained from these tests when comparing the ROC curves of Model 1 and Model 2. As Figure 5.3 shows. The curvature of the ROC for Model 1 and Model 2 indicates that soft information variables have a stronger effect on default than those of the hard information variables.

These results lead to tentative conclusions. First, soft information provides valuable input into the appraisal of credits and in predicting defaults. The results from the comparison of the hard and soft information models (Table 5.1) indicates that soft information may even be of equal importance to the value of input provided by hard information in the credit analysis of the online lending system. As the combined model with both soft and hard information has the highest predicting value, this would suggest that soft information can strengthen the risk analysis model.

5 Conclusions

We have investigated in this paper the predicting power of soft and hard information on the loan performance of the Chinese online P2P borrowers. Our results confirm the standard findings from the literature on loans extended to SMEs. Hard information is used in the Chinese P2P market, and we find that most standard hard information performs in predicting defaults in a manner that is consistent with what is known from other markets. The estimates of the impact of gender, marital status, educational level are all consistent with the literature, notwithstanding different views in the case of gender (e.g. Liao *et al.* (2015)). The length of the loan description perform very well in the estimation of the probability of default and are also consistent with Lin *et al.* (2016) but the literature is relatively new in this case.,. All three social capital proxies – weibo verification, video verification and mobile verification - are statistically significant as determinants of defaults.

However, we also find that, to our surprise, house ownership is positively related to defaults. Admittedly, this finding is not exceptional since similar results have been found in other countries. Perhaps more surprising is our result that car ownership is insignificant in our Model 1 and even negatively significant in our Model 3 as a determinant of loan defaults. It is not clear how this finding could be interpreted even though different interpretations are possible. One possible interpretation of these surprising results could be a process of adverse selection of borrowers by lenders but such an interpretation would require additional work and is rather unlikely given the Chinese system of oversight noted in the text.

It is also quite likely that loan appraisals with the help of “soft data” could be further enhanced by other and, perhaps, better proxies for social and psychological factors. Clearly, the field is open and will undoubtedly develop over time. Much “better” information already exists at various levels of business in China such as the Facebook or the Wechat data. However, it is increasingly unlikely that such data will be accessible in financial markets due to rising concerns about data privacy as exemplified by the private protection law adopted this year by the European Union and in California.

Perhaps the most interesting and somewhat surprising is the result that soft information can play an important role in credit appraisal and in predicting defaults even on its own. This can be seen from our result showing that soft information provided a powerful predictor of defaults. Moreover, even better results were obtained when we combined hard and soft information in our model. The combination of the soft and hard information had the best performance among the three. These results are consistent with experiences in the Fintech industry from other countries.¹⁴

¹⁴ While credit scores continue to be important both in the US Community Banking sector as well as by the US Fintech firms, the value of soft information in credit appraisal is increasingly recognized by both of these segments of industry. For example, several US Fintech companies indicated that they could increase intermediation by 30 percent based on their models as opposed to using standard credit scoring models. One of these companies, publicly listed Fintech Company OnDeck Capital has been able on its own to intermediate more loans than the entire US Community Finance sector. This information comes from a detailed review and analysis of the US sector and US Fintech industry carried out by a team of researchers led by Ira Lieberman on behalf of the US Treasury. We are grateful to I. Lieberman for sharing the findings with us.

Some of the policy implications of our work are evident. As our results emphasized the importance of soft information, they provide empirical evidence in support of measures to encourage greater use of soft information in addition to the hard information in online lending. Soft information should, therefore, be added to the traditional credit analysis. The importance and availability of soft information will increase with the development of technology and the information “hardening” tools. This is also in line with the expansion of credit in the big data age. If implemented, this would considerably increase the challenges for regulators. Microfinance banks and non-bank financial institutions are already regulated by local banking supervisors in the country concerned or by regional supervisors. Moreover, regulatory agencies would have to pay far more attention to lending based on the use of soft information, its quality, its dissemination and data privacy which will require a considerably different range of skills than in traditional lending. Legislative steps are very likely to be needed in order to fully reflect technological changes in Fintech industry and in financial markets.

References

- Agarwal, Sumit, and Robert Hauswald. (2010) "Distance and Private Information in Lending." *Review of Financial Studies*, 23, 2757–88.
- Agarwal, S., Ambrose, B. W., Chomsisengphet, S., & Liu, C. (2011). The role of soft information in a dynamic contract setting: Evidence from the home equity credit market. *Journal of Money, Credit and Banking*, 43(4), 633-655.
- Ashta, A., & Assadi, D. (2009). An analysis of European online micro-lending websites.
- Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 488-500.
- Barasinska, N., & Schäfer, D. (2010). Does gender affect funding success at the peer-to-peer credit markets? Evidence from the largest German lending platform. DIW Berlin Discussion Paper No. 1094
- Berger, A. N., & Udell, G. F. (2002). Small business credit availability and relationship lending: The importance of bank organizational structure. *The Economic Journal*, 112(477), F32-F53.
- Berger, A.N., Frame, W.S., Miller, N., 2005a. Credit scoring and the availability, price, and risk of small business credit. *J. Money, Credit, Banking* 37, 191–222.
- Berger, Allan, Nathan Miller, Mitchell Petersen, Raghuram Rajan, and Jeremy Stein. (2005) "Does Function Follow Organizational Form? Evidence from the Lending Practices of Large and Small Banks." *Journal of Financial Economics*, 76, 237–69.
- Bertrand, M., Karlin, D., Mullainathan, S., Shafir, E., & Zinman, J. (2005). What's psychology worth? A field experiment in the consumer credit market (No. w11892). National Bureau of Economic Research.
- Bourdieu, P. (1986). The forms of capital. *Cultural theory: An anthology*, 81-93.
- Cao Xubing (2013). Measurement and the role of social capital in online P2P lending market (Doctoral dissertation, Southwestern University of Finance and Economics).
- Dell’Ariccia, Giovanni, and Robert Marquez. (2004) "Information and Bank Credit Allocation." *Journal of Financial Economics*, 72, 185–214.
- DeYoung, R., Glennon, D., & Nigro, P. (2008). Borrower–lender distance, credit scoring, and loan performance: Evidence from informational-opaque small business borrowers. *Journal of Financial Intermediation*, 17(1), 113-143.
- Diamond, Douglas, 1984, Financial intermediation and delegated monitoring, *Review of Economic Studies* LI, 393-414.
- Duarte, J., Siegel, S., & Young, L. (2012). Trust and credit: the role of appearance in peer-to-peer lending. *Review of Financial Studies*, 25(8), 2455-2484.
- 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, 169-187.
- Everett, C. R. Group membership, relationship banking and loan default risk: the case of online social lending. *Journal of Internet Banking and Commerce*, 16, 2, 2011. Available at <www.arraydev.com/commerce/jibc/>.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, 27(8), 861-874.
- Freedman, S. M., & Jin, G. Z. (2011). Learning by Doing with Asymmetric Information: evidence from Prosper. com (No. w16855). National Bureau of Economic Research.
- García-Appendini, E. (2007). Soft information in small business lending. EFA 2007 Ljubljana Meetings Paper
- Godbillon-Camus, B., & Godlewski, C. J. (2005). Credit risk management in banks: Hard information, soft Information and manipulation. Available at SSRN: <https://ssrn.com/abstract=882027>
- Gonzalez, L., & Loureiro, Y. K. (2014). When can a photo increase credit? The impact of lender and borrower profiles on online peer-to-peer loans. *Journal of Behavioral and Experimental Finance*, 2, 44-58.
- Grunert, J., Norden, L., & Weber, M. (2005). The role of non-financial factors in internal credit ratings. *Journal of Banking & Finance*, 29(2), 509-531.
- Greiner, M.E., and Wang, H. (2009). The role of social capital in people-to-people lending marketplaces. *Proceedings of the Thirtieth International Conference on Information Systems*, Phoenix, pp. 1-17.
- Herrero-Lopez, S. (2009). Social Interactions in P2P Lending. *Proceedings of the 3rd Workshop on Social Network Mining and Analysis* (pp. 1–8).
- Horrigan, J. O. (1966). The determination of long-term credit standing with financial ratios. *Journal of Accounting Research*, 44-62.
- Jensen, M. C., & Meckling, W. H. (1992). Specific and general knowledge and organizational structure. *CONTRACT ECONOMICS*, pp. 251-274
- Jiménez, G., & Saurina, J. (2004). Collateral, type of lender and relationship banking as determinants of credit risk. *Journal of banking & Finance*, 28(9), 2191-2212.
- Klafft, M. (2008, March). Peer to peer lending: auctioning microcredits over the internet. In *Proceedings of the International Conference on Information Systems, Technology and Management*, A. Agarwal, R. Khurana, eds., IMT, Dubai.
- Lea, S. E., Webley, P., & Walker, C. M. (1995). Psychological factors in consumer debt: Money management, economic socialization, and credit use. *Journal of Economic Psychology*, 16(4), 681-701.
- Liao, L., Ji, L., & Zhang, W. (2015). Whether lending market can accurately recognize the value of diplomas? –Evidence from P2P lending platform. *Journal of Financial Research*, (3), 146-159.
- Lieberman, I. P.de Leo, T.R. Watkins and A. Kanze (2017): Microfinanced Revolution or

Footnote: The Future of Microfinance over the next 10 Years; Lehigh University, Working Paper -MF 3, October 2017.

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

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

Iyer, R., Khwaja, A. I., Luttmer, E. F., & Shue, K. (2009, August). Screening in new credit markets: Can individual lenders infer borrower creditworthiness in peer-to-peer lending? In AFA 2011 Denver meetings paper.

Miu, L., & Chen, J. (2014). The Influence of Social Capitals on Borrower's Default Risk in P2P Network Lending —A Case Study of the Prosper. *Finance Forum*, 19(3), 9-15.

Order, R., & Zorn, P. (2000). Income, location and default: Some implications for community lending. *Real Estate Economics*, 28(3), 385-404.

Petersen, M. A. (2004). Information: Hard and soft. https://scholar.google.cz/scholar?hl=en&as_sdt=0%2C5&q=+Information%3A+Hard+and+soft.&btnG=

Pope, D. G., & Sydnor, J. R. (2011). What's in a Picture? Evidence of Discrimination from Prosper. com. *Journal of Human Resources*, 46(1), 53-92.

Pötzsch, S. & R. Böhme (2010): The Role of Soft Information in Trust Building: Evidence from online Social Lending. In *International Conference on Trust and Trustworthy Computing* (pp. 381-395). Springer, Berlin, Heidelberg, June 2010.

Ravina, E. (2012). Love & loans: The effect of beauty and personal characteristics in credit markets. Available at SSRN 1107307. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1107307

Serrano-Cinca, C., Gutierrez-Nieto, B., & López-Palacios, L. (2015). Determinants of default in P2P lending. *PloS one*, 10(10), e0139427.

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

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

Stein, J. C. (2002). Information production and capital allocation: Decentralized versus hierarchical firms. *The Journal of Finance*, 57(5), 1891-1921.

Thakor, R. T. and R.C. Merton (2018): Trust in Lending; NBER, Working Paper No. 24778, Cambridge, Mass. June. <http://www.nber.org/papers/w24778>

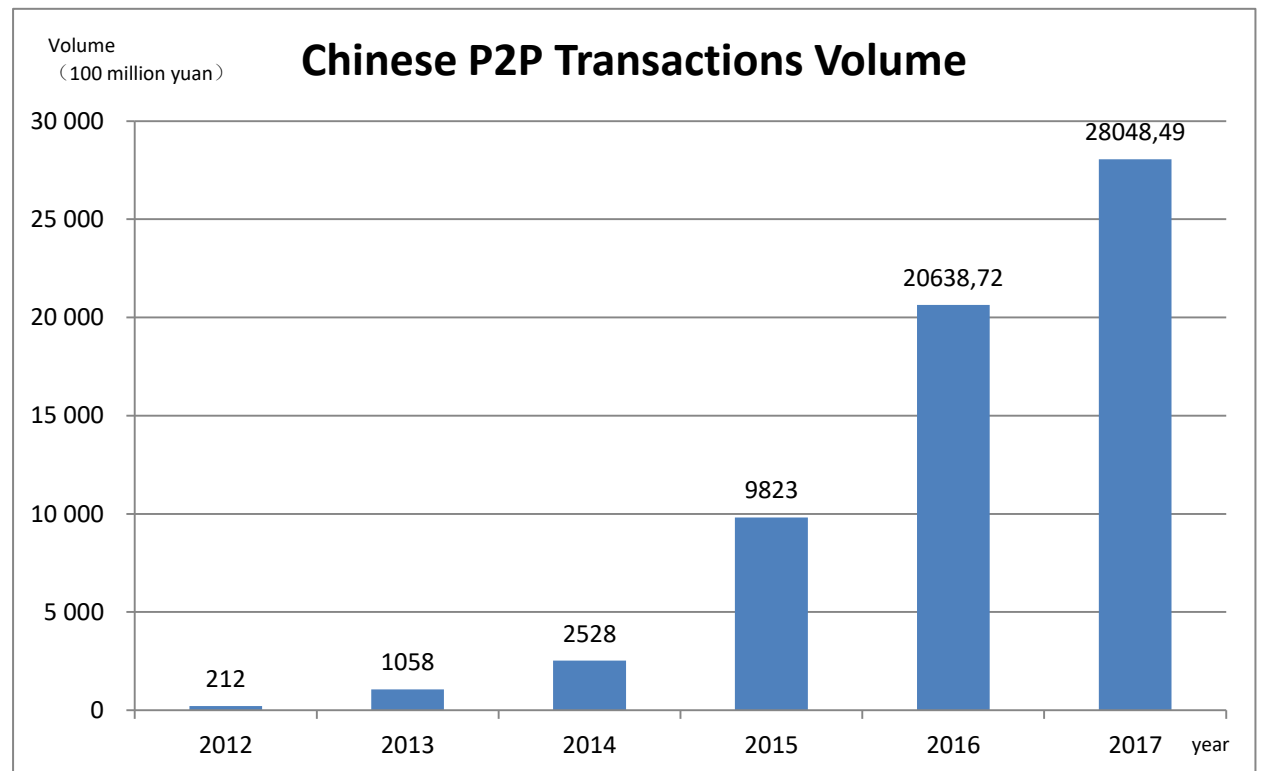
Uchida, H. (2011). What do banks evaluate when they screen borrowers? Soft information, hard information and collateral. *Journal of Financial Services Research*, 40(1-2), 29-48.

Zhong Xu, & Chuanwei Zou. (2010). The design of banks' loan approval right and incentive mechanism under hard and soft information framework—the indication to SMEs finance

problem. Journal of Financial Research, (8), 1-15.

Annex 1

Annex Table 1: Chinese P2P Market Key indicators



Annex 2

Annex Table 2 Collinearity Diagnostics

Variable	VIF	VIF	Tolerance	Squared
Income verified	1.04	1.02	0.9648	0.0352
Income	1.44	1.20	0.6932	0.3068
Car verified	1.54	1.24	0.6480	0.3520
House verified	1.65	1.28	0.6071	0.3929
Mortgage loan	1.26	1.12	0.7937	0.2063
Loan description	1.52	1.23	0.6562	0.3438
Age	1.31	1.15	0.7620	0.2380
Gender	1.03	1.02	0.9687	0.0313
Marriage	1.18	1.09	0.8478	0.1522
Educational	1.11	1.05	0.9020	0.0980
Mobile verified	1.42	1.19	0.7028	0.2972
Weibo verified	1.41	1.19	0.7068	0.2932
Video verified	1.53	1.24	0.6520	0.3480
Interest	1.10	1.05	0.9085	0.0915
Amount	2.21	1.49	0.4529	0.5471
Term	1.69	1.30	0.5928	0.4072
Mean VIF	1.40			

IES Working Paper Series

2018

1. Karel Janda, Martin Strobl: *Smoking Czechs: Modeling Tobacco Consumption and Taxation*
2. Karel Janda, Michaela Koscova: *Photovoltaics and the Slovak Electricity Market*
3. Simona Malovana, Dominika Kolcunova, Vaclav Broz: *Does Monetary Policy Influence Banks' Perception of Risks?*
4. Karolina Vozkova: *Why Did EU Banks Change Their Business Models in Last Years and What Was the Impact of Net Fee and Commission Income on Their Performance?*
5. Jan Malek, Lukas Recka, Karel Janda: *Impact of German Energiewende on Transmission Lines in the Central European Region*
6. David Svacina: *Devaluation with Exchange rate Floor in a Small Open Economy*
7. Ladislav Kristoufek: *Are the Crude Oil Markets Really Becoming More Efficient over Time? Some New Evidence*
8. Karel Janda, Zuzana Lajksnerova, Jakub Mikolasek: *A General Equilibrium Model of Optimal Alcohol Taxation in the Czech Republic*
9. Nicholas Tyack, Milan Scasny: *Estimating the Value of Crop Diversity Conservation Services Provided by the Czech National Programme for Agrobiodiversity*
10. Laure de Batz: *Financial Impact of Regulatory Sanctions on French Listed Companies*
11. Matej Opatrny: *Extent of Irrationality of the Consumer: Combining the Critical Cost Eciency and Houtman Maks Indices*
12. Mojmir Hampl, Tomas Havranek: *Foreign Capital and Domestic Productivity in the Czech Republic*
13. Miroslav Palansky: *The Value of Political Connections in the Post-Transition Period: Evidence from the Czech Republic*
14. Karel Janda: *Earnings Stability and Peer Selection for Indirect Valuation*
15. Ondrej Tobek, Martin Hronec: *Does the Source of Fundamental Data Matter?*
16. Stefan Schmelzer, Michael Miess, Milan Scasny, Vedunka Kopecna: *Modelling Electric Vehicles as an Abatement Technology in a Hybrid CGE Model*
17. Barbora Malinska, Jozef Barunik: *Volatility Term Structure Modeling Using Nelson-Siegel Model*
18. Lubomir Cingl, Vaclav Korbel: *Underlying Motivations For Rule-Violation Among Juvenile Delinquents: A Lab-in-the-Field Experiment*
19. Petr Jansky: *Estimating the Revenue Costs of Tax Treaties in Developing Countries*
20. Yao Wang, Zdenek Drabek, Zhengwei Wang: *The Predicting Power of Soft Information on Defaults in the Chinese P2P Lending Market*

All papers can be downloaded at: <http://ies.fsv.cuni.cz>



Univerzita Karlova v Praze, Fakulta sociálních věd
Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26

E-mail : ies@fsv.cuni.cz

<http://ies.fsv.cuni.cz>